## WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



### INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT GOOPERATION TREATY (PCT)

(51) International Patent Classification 5: (11) International Publication Number: WO 94/22477 A61K 39/12, 39/00, C07K 17/00, 3/00, A1 (43) International Publication Date: 13/00, 15/00, C07H 15/12, 13/00 13 October 1994 (13.10.94) (21) International Application Number: PCT/US94/03282 (81) Designated States: AU, BR, CA, CN, JP, RU, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, (22) International Filing Date: 25 March 1994 (25.03.94) IE, IT, LU, MC, NL, PT, SE). (30) Priority Data: Published 08/037,816 26 March 1993 (26.03.93) US With international search report. (60) Parent Application or Grant (63) Related by Continuation 08/037,816 (CIP) Filed on 26 March 1993 (26.03.93) (71) Applicant (for all designated States except US): PROGENICS PHARMACEUTICALS, INC. [US/US]; Old Saw Mill River Road, Tarrytown, NY 10591 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): HASEL, Karl, W. [CH/US]; 1713 Baldwin Road, Yorktown Heights, NY 10598 (US). MADDON, Paul, J. [US/US]; Apartment #25C, 60 Haven Avenue, New York, NY 10032 (US). (74) Agent: WHITE, John, P.; Cooper & Dunham, 30 Rockefeller Plaza, New York, NY 10112 (US). (54) Title: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC

**USES THEREOF** 

#### (57) Abstract

The invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein, vaccines comprising the mutant HIV-1 envelope glycoprotein, antibodies and methods of treating individuals.

> Applicants: G.P. Allaway, et al. Serial No.: 09/460,216 Filed: December 13, 1999

Exhibit 178

### FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
ΑÜ	Australia	GE	Georgia	MW	
8B	Barbados	GN	Guinea	NE	Malawi
BE	Belgium	GR	Greece	NL NL	Niger
BF	Burkina Feso	BTU	Hungary		Netherlands
BG	Bulgaria	DE .	Ireland	NO	Norway
B.J	Benin	n	Italy	NZ	New Zealand
BR	Brazil	∕ <b>ЛР</b> ·	-	PL	Poland
BY	Belans	KE	Japan	PT	Portugal
CA	Canada		Kcoya	RO	Romania
-		KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic	SD	Suden
CG	Congo		of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	K2	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	Chine	LK	Sri Lanka	170	Chad
cs	Czechoulovakia	LU	Luxembourg	TG	
CZ	Czech Republic	LV	Latvia		Togo
DE	Germany	MC	Monaco	<u>TJ</u>	Tajikistan
DK	Deemark	MD		TT	Trinidad and Tobago
ES	Spain		Republic of Moldova	UA	Ukraine
	•	MG	Madagascer	US	United States of America
<u>rı</u>	Finland	MIL	Mali	UZ	Uzhekistan
FR	France	MIN	Mongolia	VN	Vict Nam
G.A	Gabon				

# HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF

#### Background of the Invention

10

Throughout this application, various publications are referenced by Arabic numerals. Full citations for these references may be found at the end of the specification immediately preceding the claims. The disclosure of these publications is hereby incorporated by reference into this application to describe more fully the art to which this invention pertains.

The life cycle of animal viruses is characterized by a series of events that are required for the productive infection of the host cell. The initial step in the replicative cycle is the attachment of the virus to the cell surface, which attachment is mediated by the specific interaction of the viral attachment protein (VAP) to 25 receptors on the surface of the target cell. differential pattern of expression of these receptors is largely responsible for the host range and tropic properties of viruses. In addition, an effective immune response against many viruses is mediated through neutralizing 30 antibodies directed against the VAP. The interaction of the VAP with cellular receptors and the immune system therefore plays a critical role in infection and pathogenesis of viral disease.

The human immunodeficiency virus type 1 (HIV-1) infects primarily helper T lymphocytes, dendritic cells, and monocytes/macrophages--cells that express surface CD4--leading to a gradual loss of immune function. This loss of function results in the development of the human acquired

immunodeficiency syndrome (AIDS) (1). The initial phase of the HIV-1 replicative cycle involves the high-affinity interaction between the HIV-1 exterior envelope glycoprotein gp120 and cell surface CD4 ( $K_d$  approximately 4 x 10.9 M) (2). 5 Several lines of evidence demonstrate the requirement of this interaction for viral infectivity. The introduction into CD4 human cells of cDNA encoding CD4 is sufficient to render otherwise resistant cells susceptible to HIV-1 infection (3). In vivo, viral infection appears to be 10 restricted to cells expressing CD4, indicating that the cellular tropism of HIV-1 is largely determined by the pattern of cellular expression of CD4. Following the binding of HIV-1 gp120 to cell surface CD4, viral and target cell membranes fuse by a mechanism that is poorly 15 understood, resulting in the introduction of the viral capsid into the target cell cytoplasm (4).

2

Mature CD4 has a relative molecular mass (Mr) of 55 kDa and consists of an N-terminal 372-amino acid extracellular 20 domain containing four tandem immunoglobulin-like regions (V1-V4), followed by a 23-amino acid transmembrane domain and a 38-amino acid cytoplasmic segment (5, 6). experiments using truncated sCD4 proteins, it has been shown that the determinants for high-affinity binding to HIV-1 gp120 lie solely within the N-terminal immunoglobulin-like domain (V1) (7-9). Mutational analysis of V1 has defined a discrete binding site (residues 38-52) that comprises a region structurally homologous to the complementarity-determining region (CDR2) of immunoglobulin genes (9).

The production of large quantities of sCD4 has permitted a structural analysis of the two N-terminal immunoglobulinlike domains (V1V2). The structure determined at 2.3

30

3

angstrom resolution reveals that the molecule has two tightly-associated domains, each of which contains the immunoglobulin-fold connected by a continuous beta strand. The putative binding sites for monoclonal antibodies, class II major histocompatibility complex (MHC) molecules, and HIV-1 gp120, as determined by mutational analyses, map on the molecular surface (10, 11).

The HIV-1 envelope gene env encodes an envelope glycoprotein precursor, gp160, which is cleaved by cellular proteases before transport to the plasma membrane to yield gp120 and gp41. The membrane-spanning glycoprotein, gp41, is non-covalently associated with gp120, a purely extracellular glycoprotein. The mature gp120 molecule is heavily glycosylated (approximately 24 N-linked oligosaccharides), contains approximately 480 amino acid residues with 9 intrachain disulfide bonds (12), and projects from the viral membrane as a dimeric or multimeric molecule (13).

20 Mutational studies of HIV-1 gp120 have delineated important functional regions of the molecule. The regions of gp120 that interact with gp41 map primarily to the N- and Ctermini (14). The predominant strain-specific neutralizing epitope on gp120 is located in the 32-34 amino acid residue 25 third variable loop, herein referred to as the V3 loop, which resides near the center of the gp120 sequence (15). The CD4 binding site maps to discontinuous regions of gp120 that include highly conserved or invariant amino acid residues in the second, third, and fourth conserved domains 30 (the C2, C3, and C4 domains) of gp120 (16). It has been postulated that a small pocket formed by these conserved residues within gp120 could accommodate the CDR2 loop of CD4, a region defined by mutational analyses as important in interacting with gp120 (17).

HIV-1 gp120 not only mediates viral attachment to surface CD4 molecules, but also serves as the major target of antibodies which neutralize non-cell-associated virus and inhibit cell to cell viral transmission.

There are two major classifications of HIV-1-neutralizing antibodies: type-specific and group-common (15). specific neutralizing antibodies primarily recognize linear determinants in the highly variable V3 loop of gp120. These 10 antibodies act by inhibiting fusion between HIV-1 and the target cell membrane, and generally neutralize only a particular isolate of, or closely related strains of, HIV-1. Sequence variation within the V3 loop, as well as outside of this region, permits viruses to escape neutralization by 15 anti-V3 loop antibodies. In contrast, group-common neutralizing antibodies primarily recognize discontinuous or conformational epitopes in gp120, and possess the ability to neutralize a diverse range of HIV-1 isolates. These broadly neutralizing antibodies often recognize a site on gp120 which overlaps the highly conserved CD4-binding site, and thus inhibits gp120-CD4 binding.

A structural relationship has been demonstrated between the V3 loop and the C4 region of gp120 which region constitutes 25 both part of the CD4 binding site and part of the conserved neutralization epitopes. It was observed that deleting the V3 loop resulted in significantly increased binding of a panel of broadly neutralizing hMoAbs (neutralizing human monoclonal antibodies) to the CD4 binding site (18).

30

20

A major goal in AIDS vaccine development is to develop a vaccine able to protect a subject against the numerous genetic variants of HIV-1 that infect humans. cell-mediated immune responses might serve to control 35 infection in HIV-1-infected individuals, several lines of

PCT/US94/03282 WO 94/22477

evidence demonstrate that protection against infection is mainly mediated by neutralizing antibodies directed against Early experiments showed that immunization of chimpanzees with recombinant gpl20 induced a protective 5 immune response against challenge with the homologous HIV-1 strain (17). This protection correlated with the presence of high-titer neutralizing antibodies against the V3 loop of In addition, passive immunization of chimpanzees with a V3-loop neutralizing monoclonal antibody resulted in 10 protection against challenge with the homologous HIV-1 Although protection against challenge was strain (19). demonstrated in these two experiments, recent studies have questioned the clinical relevance of these findings. example, these neutralizing antibodies recognize the V3 loop 15 determinants of a single strain, and not conserved or Thus, these antibodies lack the discontinuous epitopes. ability to neutralize the broad spectrum of HIV-1 strains present in an HIV-1 population. Furthermore, the challenge virus was the homologous HIV-1 laboratory adapted LAI (HTLV-IIIB) strain and not one of the primary isolates that contain considerable gp120 sequence heterogeneity. these experiments showed that gp120 subunit vaccination induces an immune response effective against only the homogeneous HIV-1 strain used as an antigen, it is unlikely 25 that the vaccination regimens used in these studies would be useful in humans.

Individuals infected by HIV-1 typically develop antibodies that neutralize the virus in vitro, and neutralization 30 titers decrease with disease progression (19). Analysis of sera from HIV-1-infected humans indicates that type-specific neutralizing antibodies appear early in infection. Later in the course of infection, a more broadly neutralizing antibody response develops. However this antibody response is of significantly lower titer and/or affinity.

35

6

Fractionation studies of HIV-1 antibody-positive human sera reveal that the type-specific neutralizing activity is primarily directed against linear determinants in the V3 loop of gp120 (20). There was no correlation found among antibodies between the ability to neutralize divergent HIV-1 isolates and reactivity to the V3 loop of these isolates. In contrast, the broadly neutralizing antibodies present in HIV-1 antibody-positive human sera primarly recognize discontinuous epitopes in gp120 which overlap the CD4-binding site and block gp120-CD4 binding. In other words, the broadly neutralizing activity of neutralizing antibodies is not merely the result of additive anti-V3 loop reactivities against diverse HIV-1 isolates which appear during infection.

15

Recently, several groups have generated human monoclonal antibodies (hMoAbs) derived from HTV-1 infected individuals which possess type-specific or group-common neutralizing activities (17). The type-specific neutralizing hMoAbs were found to recognize linear determinants in the V3 loop of gp120. In contrast, the group-common neutralizing hMoAbs generally recognize discontinuous epitopes which overlap the CD4-binding site and block gp120-CD4 binding.

The V3 loop is a highly immunodominant region of gp120 which partially interacts with the CD4-binding region. The presence of the V3 loop region on gp120 may skew the humoral immune response away from producing antibodies which specifically bind to the CD4-binding domain of gp120.

Furthermore, the advantages of removing the V3 loop to expose the CD4-binding domain of gp120 to the immune system would be countered by the fact that the exposed CD4-binding site would still have a high affinity for cell surface CD4. In other words, a mutant gp120 protein missing only the V3 loop would quickly bind to CD4+ cells and would thus be

7

hampered in generating an immune response against the exposed CD4-binding site.

The subject invention provides a mutant HIV-1 gp120 envelope glycoprotein which overcomes both the problems of V3 loop immunodominance and of the high affinity to CD4. The subject invention further provides vaccines comprising the mutant HIV-1 gp120 envelope glycoprotein, antibodies which specifically bind to the CD4-binding site of HIV-1 gp120 envelope glycoprotein, pharmaceutical compositions comprising these antibodies, and methods of using these vaccines and compositions to treat or prevent HIV-1 infection.

15

#### Summary of the Invention

The subject invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(W->X)</sub> point mutation, wherein X is an amino acid residue other than tryptophan. In the preferred embodiment, X is a valine residue.

- 10 In one embodiment, the nucleic acid molecule is a DNA molecule. The DNA molecule may be a plasmid. In one embodiment, the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
- 15 In one embodiment, the C4 domain is an HIV-1<sub>LAI</sub> gp120 envelope glycoprotein C4 domain. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein.
- In another embodiment, the C4 domain is an HIV-1 $_{\rm R-FL}$  gp120 envelope glycoprotein C4 domain. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1 $_{\rm R-FL}$  gp120 envelope glycoprotein.
- 25 The subject invention also provides the mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of the subject invention.
- The subject invention further provides a vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.

The subject invention further provides a method of treating

9

an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of the subject invention, thereby treating the HIV-1-infected subject.

- 5 The subject invention further provides a vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

The subject invention further provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

25 partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of the subject invention, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein. In the preferred embodiment, the subject is a human.

The subject invention further provides the partially purified antibodies produced by the method of the subject invention.

10

- 5 The subject invention further provides a pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.
- The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

The subject invention further provides a composition which comprises a prophylactically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises administering to the HIV-1-exposed subject a dose of the composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-exposed subject, thereby reducing the likelihood of the subject's becoming infected with HIV-1.

In one embodiment, the subject is a medical practitioner. In another embodiment, the subject is a newborn infant.

Finally, the subject invention provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1. In one embodiment, the subject is a medical practitioner.

#### Brief Description of the Figures

#### Figure 1

depicting the boundaries of the five constant domains (C1-C5) and the five variable domains (V1-V5). The amino acid residue numbering above the box begins at the initiator methionine found at the beginning of the signal sequence (S) and is approximated based on a consensus of all known HIV-1 gp120 amino acid sequences. Also shown are the C4 domain amino acid sequences of HIV-1 strains LAI and JR-FL. Above the C4 domain sequences are indicated two mutations that reduce gp120 binding to cell surface CD4; tryptophan to valine and aspartate to alanine.

15

#### Figure 2

#### Figure 3

cmv MIE promoter fused to tPA-gpl20<sub>LAI</sub>. The nucleotide sequence of the CMV MIE promoter/enhancer region is shown fused to the HIV-1<sub>LAI</sub> gpl20 gene that contains the tPA signal sequence. The numbering of nucleotide sequence begins with the HincII site and the numbering of the amino acid sequence begins with the first methionine found in the tPA signal sequence. The tPA signal sequence is fused in-frame to Thr31

13

of gp120, the first amino acid found in mature gp120. The signal sequence is shown in bold as are various landmark restriction sites used for cloning as discussed in the text. The locations of Exon A, Intron A, Exon B and the transcription start site and the signal cleavage site are indicated.

#### Figure 4

Transient expression of gp120. Autoradiograph of <sup>35</sup>S-labeled supernatants from COS cell transfectants, immunoprecipitated with a CD4-immunoglobulin-Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. The plasmids used for transfection were: Lane 1: Mock transfected cells; lane 2: a vector encoding a CD4-immunoglobulin chimera as a positive transfection control; lane 3: PPI4-tPA-gp120<sub>LAI</sub>; and lane 4: PPI4-tPA-gp120<sub>IR-FL</sub>. Positions of molecular weight markers are indicated.

#### Figure 5

20 <u>Determination of gp120 concentration by ELISA</u>. Panel A: Concentrations of gp120 in media of CHO cell lines, stably transfected with PPI4-tPA-gp120<sub>LAI</sub>, determined by ELISA. Panel B: A standard curve was established using known amounts of gp120.

25

#### Figure 6

Expression of gp120 in stably transfected CHO cells.

Autoradiograph of <sup>35</sup>S-labeled supernatants from stable CHO cell lines, immunoprecipitated with a CD4-immunoglobulin
Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. Lane 1: clone 9; lane 2: clone 13; lane 3: clone 6; lane 4: Clone 5. Positions of molecular weight markers are indicated.

14

#### Figure 7

tPA-gp120<sub>JR-FL</sub>. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>JR-FL</sub> gp120 is shown. The NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Val<sub>36</sub> is indicated.

#### Figure 8

tPA-gp120<sub>LAI</sub>-V3<sup>(4)</sup>. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>LAI</sub> gp120 with the V3 loop deleted and replaced with the pentapeptide TGAGH is shown. The V3 loop replacement and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Thr<sub>36</sub> is indicated.

#### Figure 9

tPA-gp120<sub>TR-FL</sub>-V3<sup>(\*)</sup>. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>JR-FL</sub> gp120 with the V3 loop deleted and replaced with the pentapeptide TGAGH is shown. The V3 loop replacement and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Val<sub>36</sub> is indicated.

25

#### Figure 10

tPA-qp120<sub>LAI</sub>-V3<sup>(\*)</sup>-CD4<sup>(\*)</sup>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>LAI</sub> gp120, with the V3 loop deleted and replaced with the pentapeptide TGAGH, and Trp403 mutated to Val. The mutations and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg35 and Thr36 is indicated.

#### Figure 11

tPA-gp120<sub>JR-FL</sub>-V3<sup>(\*)</sup>-CD4<sup>(\*)</sup>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>JR-FL</sub> gp120, with the V3 loop deleted and replaced with the pentapeptide TGAGH, and Trp<sub>3%</sub> mutated to Val. The mutations and the Narl and Notl restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Val<sub>36</sub> is indicated.

10

#### Figure 12

tPA-gpl20<sub>LAI</sub>-CD4<sup>()</sup>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>LAI</sub> gpl20. The Trp<sub>437</sub> to Val CD4 binding mutation, the Narl and NotI restriction endonuclease sites used for cloning, and the predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Thr<sub>36</sub> are shown in bold.

#### Figure 13

tPA-gp120<sub>R-FL</sub>-CD4<sup>()</sup>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>R-FL</sub> gp120. The Trp<sub>424</sub> to Val CD4 binding mutation, the NarI and NotI restriction endonuclease sites used for cloning and the predicted cleavage by signal peptidase between Arg<sub>33</sub> and Val<sub>36</sub> are shown in bold.

#### Figure 14

Expression of qp120 in stably transfected CHO cells.

Autoradiograph of super <sup>35</sup>S-labeled supernatants from stable CHO cell lines, immunoprecipitated with MoAb F105-Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. Panel A: Lane 1: tPA-gp120<sub>LAI</sub> CHO cells; lane 2: tPA-gp120<sub>LAI</sub>-V3<sup>(-)</sup> CHO cells; lane 3: tPA-gp120<sub>LAI</sub>-V3<sup>(-)</sup>-CD4<sup>(-)</sup> CHO cells. Panel B: Lane 1: tPA-gp120<sub>IR-FL</sub> CHO cells; lane 2: tPA-gp120<sub>IR-FL</sub>-V3<sup>(-)</sup>

CHO cells; lane 3: tPA-gp120<sub>JR-FL</sub>-V3<sup>(-)</sup>-CD4<sup>(-)</sup> CHO cells. Positions of molecular weight markers are indicated.

#### Figure 15

5 Purified ap120 proteins.

Silver stained 10% SDS-PAGE gel with a sample of purified gp120 proteins. Panel A: Lane 1: tPA-gp120<sub>LAI</sub> CHO cells; lane 2: tPA-gp120<sub>LAI</sub>-V3<sup>(·)</sup> CHO cells; lane 3: tPA-gp120<sub>LAI</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup> CHO cells. Panel B: Lane 1: tPA-gp120<sub>IR-FL</sub> CHO cells; lane 2: tPA-gp120<sub>IR-FL</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup> CHO cells. Positions of molecular weight markers are indicated.

#### Figure 16

Analysis of binding of recombinant mutant gp120 to cell surface human CD4 by FACS.

Plate 1. DG44 cells, a subclone of CHO cells which lack expression of the human CD4 protein, were used as control. Increasing concentrations of HIV-1 gp120<sub>IAI</sub> did not show an compared specific fluoresence when in Plate 2. DG44 #3 cells are a CHO cell line 20 background. transfected with the cDNA clone encoding the human CD4 protein. Increasing concentrations of HIV-1 gp120LAI show a dramatic increase (or shift) in fluoresence. Similar to Plate 2 but the HIV-1 gp120LAI-V3() protein was 25 added. Again a large shift indicating binding to the DG44 #3 cells was seen. Plate 4. DG44 #3 cells were incubated with either HIV-1 gp120<sub>LAI</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup> protein or MoAb OKT4A an antibody with high affinity for human CD4. Only OKT4A bound to the cells.

#### Detailed Description of the Invention

The plasmids designated PPI4-tPA-gp120<sub>LAI</sub> and PPI4-tPA-gp120<sub>JR</sub> fl were deposited pursuant to, and in satisfaction of, the requirements of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland 20852 under ATCC Accession Nos. 75431 and 75432, respectively. The plasmids PPI4-tPA-gp120<sub>LAI</sub> and PPI4-tPA-gp120<sub>JR-FL</sub> were deposited with the ATCC on March 12, 1993.

The subject invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(W->X)</sub> point mutation, wherein X is an amino acid residue other than tryptophan. In the preferred embodiment, X is a valine residue.

- In one embodiment, the nucleic acid molecule is a DNA molecule. The DNA molecule may be a plasmid. In one embodiment, the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
- 25 The V3 loop of HIV-1 gp120 envelope glycoprotein is shown in Figure 1. The V3 loop is demarcated by cysteine residues at both its N- and C-termini. As used herein, a V3 loop deletion means a deletion of one or more amino acid residues between the terminal cysteine residues, with the proviso 30 that there must be three or more amino acid residues situated between the two terminal cysteine residues in a V3 loop deletion. These three or more amino acid residues may either be residues originally present in the V3 loop, or exogenous residues. For example, as shown in the

Experimental Details section <u>infra</u>, the pentapeptide TGAGH is situated between the two terminal cysteine residues. Variations in the size of the V3 loop deletion illustrated herein are tolerable without affecting the overall structure of the mutant HIV-1 gp120 envelope glycoprotein, as is well known to those skilled in the art.

As used herein, "C4 domain" means the HIV-1 gp120 envelope glycoprotein C4 domain having the following consensus 10 sequence:

 $\begin{array}{l} X_{1}X_{2}X_{3}CX_{4}IX_{5}X_{6}X_{7}X_{8}X_{9}X_{10}WX_{11}X_{12}X_{13}X_{14}X_{15}AX_{16}YX_{17}X_{18}-\\ PX_{19}X_{20}X_{21}X_{22}X_{24}X_{25}X_{26}SX_{27}X_{28}TGX_{29}X_{30}X_{31}X_{32}RX_{33}GX_{34}, \end{array}$ 

15 wherein X<sub>1</sub> = T, I, V, K or R; X<sub>2</sub> = L, I or H; X<sub>3</sub> = P, Q, L or
T; X<sub>4</sub> = R, K or G; X<sub>5</sub> = K or E; X<sub>6</sub> = Q or E; X<sub>7</sub> = F, I or V;
X<sub>8</sub> = I, V or M; X<sub>9</sub> = N, R or K; X<sub>10</sub> = M, R, L or T; X<sub>11</sub> = Q, R
or V; X<sub>12</sub> = E, K, G, R, V or A; X<sub>13</sub> = V, T, A or G; X<sub>14</sub> = G or
E; X<sub>15</sub> = K, R, E, or Q; X<sub>16</sub> = M, V, I or L; X<sub>17</sub> = A, T or D; X<sub>18</sub>
20 = P or L; X<sub>19</sub> = I or F; X<sub>20</sub> = S, R, G, K, N, A, E or Q; X<sub>21</sub> =
G or R; X<sub>22</sub> = Q, L, P, N, K, V, T, E or I; X<sub>23</sub> = I, V or L; X<sub>24</sub>
= R, K, S, N, G, I, T, E or I; X<sub>25</sub> = C or R; X<sub>26</sub> = S, L, I, T,
P, E, V, K, D or N; X<sub>27</sub> = N, K or L; X<sub>28</sub> = I or V; X<sub>29</sub> = L, P
or I; X<sub>30</sub> = L or I; X<sub>31</sub> = L or I; X<sub>32</sub> = T, A, I, V or E; X<sub>33</sub> =
25 D or E; X<sub>34</sub> = G or V.

The C4 domain consensus sequence is based on existing C4 domain sequence information from various HIV-1 strains, and thus is not necessarily an exhaustive consensus sequence. The conserved tryptophan residue shown in bold after residue  $X_{10}$  is the only conserved tryptophan residue in the C4 domain. As used herein, a C4 domain $_{(W->X)}$  point mutation is a mutation of the above-identified conserved C4 domain tryptophan residue to an amino acid residue other than

19

tryptophan. For example, a C4 domain $_{(W->V)}$  point mutation is a mutation of the conserved C4 domain tryptophan residue to a valine residue.

In one embodiment, the C4 domain is an HIV-1<sub>LAI</sub> gp120 envelope glycoprotein C4 domain. The sequence of the HIV-1<sub>LAI</sub> gp120 C4 domain is: TLPCRIKQFINMWQEVGKAMYAPPISGQIRCS-SNITGLLLTRDGG. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein.

10

15

In another embodiment, the C4 domain is an  $\text{HIV-1}_{\text{IR-FL}}$  gp120 envelope glycoprotein C4 domain. The sequence of the  $\text{HIV-1}_{\text{IR}}$  gp120 C4 domain is: TLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS-SNITGLLLTRDGG. The mutant HTV-1 gp120 envelope glycoprotein may be a mutant HTV-1<sub>IR-FL</sub> gp120 envelope glycoprotein.

HIV-1<sub>LAI</sub> is a laboratory-adapted strain that is tropic for phytohemagglutinin (PHA)-stimulated peripheral lymphocytes (PBLs) and immortalized human T-cell lines. 20 contrast,  $\text{HIV-1}_{\text{IR-FL}}$  was isolated from brain tissue taken at autopsy that was co-cultured with lectin-activated normal human PBLs. HIV-1<sub>IR-FL</sub> is tropic for PHA-stimulated PBLs and blood-derived macrophages but will not replicate transformed T-cell lines. Mutant HIV-1 gp120 envelope 25 glycoproteins derived from a clinical isolate of HIV-1 such as JR-FL may possess new or different epitopes compared to the laboratory-adapted HIV-1 strains that are beneficial for successful vaccination. Although only the HIV-1 and HIV- $\mathbf{1}_{\text{R-FL}}$  strains are used herein to generate the mutant HIV-1 30 gp120 envelope glycoproteins of the subject invention, other HIV-1 strain could be substituted in their place as is well known to those skilled in the art.

The V1 and V2 variable regions of gp120 are unnecessary for

CD4 binding (21). Therefore the mutant HIV-1 gp120 envelope glycoprotein of this invention can either include or exclude the V1 and V2 variable regions.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(Asp->X)</sub> point mutation, wherein the aspartate residue is between amino acid residues  $X_{15}$  and  $X_{16}$  in the C4 consensus sequence, and X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is an alanine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(Gh->X)</sub> point mutation, wherein the glutamate residue is between amino acid residues X<sub>15</sub> and X<sub>16</sub> in the C4 consensus sequence, and X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is an alanine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain<sub>(ap)778->X)</sub> point mutation, wherein X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is a lysine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>R-FL</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain<sub>(asp369->X)</sub> point mutation, wherein X is an amino acid residue other than aspartate or glutamate. In the preferred

PCT/US94/03282

embodiment, X is a lysine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain<sub>(glu380->X)</sub> point mutation, wherein X is an amino acid residue other than glutamate. In the preferred embodiment, X is a glutamine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>JR-FL</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain<sub>(gh371->X)</sub> point mutation, wherein X is an amino acid residue other than glutamate. In the preferred embodiment, X is a glutamine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C2 domain<sub>(th/267->X)</sub> point mutation, wherein X is an amino acid residue other than threonine. In the preferred embodiment, X is an arginine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HTV-1<sub>IR-FL</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C2 domain<sub>(th/260->X)</sub> point mutation, wherein X is an amino acid residue other than threonine. In the preferred embodiment, X is an arginine residue.

30

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising (a) a V3 loop deletion, or (b) a one of the C2, C3 or C4 domain point mutations

discussed supra.

The point mutations in the recombinant nucleic acid molecules described <u>supra</u> are selected based on their ability to reduce the affinity of the mutant gp120 glycoprotein encoded thereby for CD4. As used herein, the term "reduce the affinity" means to reduce the affinity by at least two-fold.

One skilled in the art would know how to make recombinant nucleic acid molecules which encode mutant HIV-1 gp120 envelope glycoproteins comprising a V3 loop deletion and the specific C2, C3 or C4 domain point mutations corresponding to those mutations exemplified in the HIV-1<sub>R-FL</sub> and HIV-1<sub>LA</sub> strains, supra. Furthermore, one skilled in the art would know how to use these recombinant nucleic acid molecules to obtain the proteins encoded thereby, and practice the therapeutic and prophylactic methods of using same, as described herein for the recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(W->X)</sub> point mutation.

The subject invention also provides the mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of the subject invention.

In accordance with the invention, numerous vector systems for expression of the mutant HIV-1 gp120 envelope glycoprotein may be employed. For example, one class of vectors utilizes DNA elements which are derived from animal viruses such as bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (RSV, MMTV or MoMLV), Semliki Forest virus or SV40 virus.

Additionally, cells which have stably integrated the DNA into their chromosomes may be selected by introducing one or more markers which allow for the selection of transfected host cells. The marker may provide, for example, prototropy auxotrophic host, biocide resistance, antibiotics) or resistance to heavy metals such as copper or the like. The selectable marker gene can be either directly linked to the DNA sequences to be expressed, or introduced into the same cell by cotransformation. Additional elements 10 may also be needed for optimal synthesis of mRNA. signals, include splice as well may promoters, enhancers, and termination transcriptional The cDNA expression vectors incorporating such elements include those described by Okayama (22).

23

15

The vectors used in the subject invention are designed to express high levels of mutant HIV-1 gp120 envelope glycoproteins in cultured eukaryotic cells as well as efficiently secrete these proteins into the culture medium.

The targeting of the mutant HIV-1 gp120 envelope glycoproteins into the culture medium is accomplished by fusing in-frame to the mature N-terminus of the mutant HIV-1 gp120 envelope glycoprotein the tissue plasminogen activator (tPA) prepro-signal sequence.

25

The mutant HIV-1 gp120 envelope glycoprotein may be produced by a) transfecting a mammalian cell with an expression vector for producing mutant HIV-1 gp120 envelope glycoprotein; b) culturing the resulting transfected mammalian cell under conditions such that mutant HIV-1 gp120 envelope glycoprotein is produced; and c) recovering the mutant HIV-1 gp120 envelope glycoprotein so produced.

Once the expression vector or DNA sequence containing the constructs has been prepared for expression, the expression

24

vectors may be transfected or introduced into an appropriate mammalian cell host. Various techniques may be employed to achieve this, such as, for example, protoplast fusion, calcium phosphate precipitation, electroporation or other conventional techniques. In the case of protoplast fusion, the cells are grown in media and screened for the appropriate activity. Expression of the gene encoding a mutant HIV-1 gp120 envelope glycoprotein results in production of the mutant glycoprotein.

10

15

Methods and conditions for culturing the resulting transfected cells and for recovering the mutant HIV-1 gp120 envelope glycoprotein so produced are well known to those skilled in the art, and may be varied or optimized depending upon the specific expression vector and mammalian host cell employed.

In accordance with the claimed invention, the preferred host cells for expressing the mutant HTV-1 gp120 envelope glycoprotein of this invention are mammalian cell lines. Mammalian cell lines include, for example, monkey kidney CV1 line transformed by SV40 (COS-7); human embryonic kidney line 293; baby hamster kidney cells (BHK); Chinese hamster ovary-cells-DHFR (CHO); Chinese hamster ovary-cells DHFR (DXB11); monkey kidney cells (CV1); African green monkey kidney cells (VERO-76); human cervical carcinoma cells (HELA); canine kidney cells (MDCK); human lung cells (W138); human liver cells (Hep G2); mouse mammary tumor (MMT 060562); mouse cell line (C127); and myeloma cell lines.

30

Other eukaryotic expression systems utilizing non-mammalian vector/cell line combinations can be used to produce the mutant HIV-1 gp120 envelope glycoproteins. These include, but are not limited to, baculovirus vector/insect cell

expression systems and yeast shuttle vector/yeast cell expression systems.

Methods and conditions for purifying mutant HIV-1 gp120 envelope glycoproteins from the culture media are provided in the invention, but it should be recognized that these procedures can be varied or optimized as is well known to those skilled in the art.

- 10 The subject invention further provides a vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- 15 A therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein may be determined according to methods well known to those skilled in the art.

As used herein, adjuvants include, but are not limited to, alum, Freund's incomplete adjuvant (FIA), Saponin, Quil A, Monophosphoryl lipid A (MPL), and nonionic block copolymers (SAF) such as L-121 (Pluronic; Syntex SAF). In the preferred embodiment, the adjuvant is alum, especially in the form of a thixotropic, viscous, and homogeneous aluminum hydroxide gel. The vaccine of the subject invention may be administered as an oil in water emulsion. Methods of combining adjuvants with antigens are well known to those skilled in the art.

- 30 The subject invention further provides a method of treating an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of the subject invention, thereby treating the HIV-1-infected subject.
- 35 As used herein, treating an HIV-1-infected subject with the

vaccine of the subject invention means reducing in the subject either the population of HIV-1 or HIV-1-infected cells, or ameliorating the progression of an HIV-1-related disorder in the subject.

5

As used herein, an "HIV-infected subject" means an individual having at least one of his own cells invaded by HIV-1.

- 10 As used herein, "immunizing" means administering a primary dose of the vaccine to a subject, followed after a suitable period of time by one or more subsequent administrations of the vaccine, so as to generate in the subject an immune response against the CD4-binding region of the mutant HIV-1 gp120 envelope glycoprotein in the vaccine. A suitable period of time between administrations of the vaccine may
- period of time between administrations of the vaccine may readily be determined by one skilled in the art, and is usually in the order of several weeks to months.
- 20 In the preferred embodiment, the dose of vaccine administered is an amount sufficient to deliver to the subject between 10ug and 1mg of the mutant HIV-1 gp120 envelope glycoprotein.
- The subject invention further provides a vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- 30 A prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein may be determined according to methods well known to those skilled in the art.
- The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming

PCT/US94/03282

infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

5

As used herein, the subject's becoming infected with HIV-1 means the invasion of the subject's own cells by HIV-1.

As used herein, reducing the likelihood of a subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 by at least two-fold. For example, if a subject has a 1% chance of becoming infected with HIV-1, a two-fold reduction in the likelihood of the subject's becoming infected with HIV-1 would result in the subject's having a 0.5% chance of becoming infected with HIV-1. In the preferred embodiment of this invention, reducing the likelihood of the subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 by at least ten-fold.

As used herein, an HIV-1-exposed subject is a subject who has HIV-1 present in his body, but has not yet become HIV-1-infected.

25

The subject invention further provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

As used herein, a non-HIV-1-exposed subject is a subject who does not have HIV-1 present in his body.

The subject invention further provides a method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of the subject invention, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein. In the preferred embodiment, the subject is a human.

As used herein, partially purified antibodies means a composition which comprises antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, and consists of fewer protein impurities than does the serum from which the anti-CD4-binding domain antibodies are derived. A protein impurity means a protein other than the anti-CD4-binding domain antibodies. For example, the partially purified antibodies might be an IgG preparation.

Methods of recovering serum from a subject are well known to those skilled in the art. Methods of partially purifying antibodies are also well known to those skilled in the art, and include, by way of example, filtration, ion exchange chromatography, and precipitation.

In one embodiment, the partially purified antibodies comprise an immune globulin (IG) preparation. IG can be purified from serum by a two-step process. Initially, serum is fractionated by the cold ethanol method of Cohn, et al. (29). Cohn Fraction II has as its main protein component IgG immunoglobulin present as monomers, dimers and aggregates. Fraction II is then purified to produce IVIG

(immune globulin intravenous) using a variety of purification methods which include, for example, ion exchange, DEAE chromatography, acid pH 4.25 diafiltration, PEG precipitation or Pepsin treatment. The final product is stabilized (e.g., glucose + NaCl) and the final IgG concentration is fixed at between about 3% and about 6%.

The subject invention further provides the partially purified antibodies produced by the method of the subject invention.

The subject invention further provides a pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

A therapeutically effective amount of the partially purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

20

Pharmaceutically acceptable carriers are well known to those skilled in the art and include, but are not limited to, 0.01-0.1M and preferably 0.05M phosphate buffer or 0.8% Additionally, such pharmaceutically acceptable saline. 25 carriers may be aqueous or non-aqueous suspensions, and emulsions. Examples of non-aqueous propylene glycol, polyethylene solvents are vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include solutions, emulsions alcoholic/aqueous 30 water, or including saline and buffered suspensions, Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid 35 and nutrient replenishers, electrolyte replenishers such as

those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

30

5

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

As used herein, administering may be effected or performed using any of the various methods known to those skilled in the art. The administering may comprise administering intravenously. The administering may also comprise administering intramuscularly. The administering may further comprise administering subcutaneously.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

30

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected subject, thereby treating the HIV-1-

31

infected subject.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

The subject invention further provides a composition which comprises a prophylactically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

A prophylactically effective amount of the partially 20 purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

The subject invention further provides a method of reducing
the likelihood of an HIV-1-exposed subject's becoming
infected with HIV-1, which comprises administering to the
HIV-1-exposed subject a dose of the composition of the
subject invention effective to reduce the population of HIV1 in the HIV-1-exposed subject, thereby reducing the
likelihood of the subject's becoming infected with HIV-1.

In one embodiment, the subject is a medical practitioner.

The medical practitioner may be a medical practitioner exposed to an HIV-1-containing bodily fluid. As used herein,

the term "medical practitioner" includes, but is in no way

PCT/US94/03282

limited to, doctors, dentists, surgeons, nurses, medical laboratory assistants, and students in health care programs.

In another embodiment, the subject is a newborn infant. The newborn infant may be a newborn infant born to an HIV-1-infected mother.

The dose of the composition of the subject invention effective to reduce the population of HIV-1 in the HIV-110 exposed subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

The vaccines and pharmaceutical compositions of the subject invention may also ameliorate the progression of an HIV-1-related disorder in a subject to whom the vaccines or pharmaceutical compositions were administered while the subject was either non-HIV-1-exposed or HIV-1-exposed, but not yet HIV-1-infected.

25

20

Finally, the subject invention provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1. In one embodiment, the

subject is a medical practitioner.

An incident wherein there is an increased risk of exposure to HIV-1 includes, for example, receiving a blood transfusion, sexual contact with an HIV-1-infected individual, and performing a HIV-1-containing bodily fluid-exposing medical procedure.

As used herein, "immediately prior to the incident" means of within one month of the incident. In the preferred embodiment, "immediately prior to the incident" means within one day of the incident.

The dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100mg/kg and 2g/kg of protein if administered intravenously.

invention is a method embodiment this of 25 One substantially reducing the likelihood of a non-infected medical practitioner's becoming infected with HIV-1 during a bodily fluid-exposing medical procedure involving a patient, which comprises administering to the patient during a suitable time period an amount of the composition of the 30 subject invention effective to substantially reduce the likelihood of the non-infected medical practitioner's becoming infected with HIV-1 by virtue of contact with the patient's bodily fluid during the medical procedure.

10

15

25

As used herein, a bodily fluid is any fluid which is present in the human body and is capable of containing infectious HIV-1 in an HIV-1-infected patient. Bodily fluids include, but are not limited to, saliva, cerebrospinal fluid, tears, vaginal secretions, urine, alveolar fluid, synovial fluid and pleural fluid.

Another embodiment of this invention is a method of substantially reducing the likelihood of a non-HIV-1-infected newborn infant's becoming infected with HIV-1 prior to or during birth from an HIV-1-infected mother, which comprises administering to the mother prior to birth an amount of the composition of the subject invention effective to substantially reduce the likelihood of the non-HIV-1-infected newborn infant's becoming infected with HIV-1 by virtue of contact with the patient's bodily fluid.

In order to facilitate an understanding of the Experimental Details section which follows, certain frequently occurring methods and/or terms are best described in Maniatis et al. (23).

This invention will be better understood by reference to the Experimental Details which follow, but those skilled in the art will readily appreciate that the specific experiments detailed are only illustrative of the invention as described more fully in the claims which follow thereafter.

25

## Experimental Details

## Nomenclature

As used herein, V3(-) indicates a V3 loop deletion from HIV-1 5 gp120 envelope glycoprotein. As used herein, CD4 indicates a point mutation in the C4 domain of HIV-1 gp120 envelope glycoprotein which mutation inhibits CD4 binding to the mutant HIV-1 gp120 envelope glycoprotein. The structure of HIV-1 gp120 envelope glycoprotein is shown in Figure 1.

#### Materials and Methods 10

Construction of PPI4-tPA-gp120 expression vector. An expression vector was constructed that consisted of the immediate early (CMV major cytomegalovirus promoter/enhancer linked to the HIV-1LAI env gene, which gene had its signal sequence replaced by the tPA signal sequence. The CMV MIE promoter/enhancer sequences were derived from pSVCC1 (24) consisting of 1580 base pairs of contiguous DNA that is immediately 5' to the initiator ATG. In sequential order, the functional domains of the CMV promoter are: the promoter/enhancer region; a transcriptional initiator site; exon A (a non-coding exon); intron A; and 17 nucleotides of exon B (non-coding sequences). The viral promoter sequences were ligated to a gene construct consisting of the nucleotide sequences encoding amino acids -35 to -1 of human tPA (25) fused in-frame to HIV-1<sub>LAI</sub> env amino acids 31 through 515, ending with a TGA stop codon. The construction was performed in two parts. The majority of the CMV promoter could be isolated as a 1560 bp Hinc II/Pst I fragment which 30 was ligated to a Pst I/Not I 1590 bp DNA fragment that contained the remainder of the CMV promoter, the initiator ATG, the tPA signal sequence and the mature HIV-1LAI env protein coding sequence.

The latter fragment was assembled using the polymerase chain Primer 1 (GATCCTGCAGTCACCGTCCTTGACAreaction as follows. CGATGGATGCAATGAAGAGA) and primer 2 (AAGTCTTCTCCTCGGTCTTGT-CTTTTTAACACCCAG) were used to amplify the nucleic acid 5 sequences encoding the tPA signal sequence amino acids -35 to -1 from plasmid pMAM neo-s (Clonetech), thus producing a 150 bp fragment. A second 1440 bp DNA fragment was amplified (TTCAGAAGAGGAGCCAGAACAGAAAAATTGTGGGTC). primer 3 primer 4 (GGAAAAAGCGGCCGCTCATTTTCTCTCTGCACCACTC), and pENV The PCR fragments were pooled, as a template. 10 desalted, and excess primer removed by ultrafiltration through a centricon-100 unit (Amicon). An aliquot of the pooled material was then subjected to a second round of amplification in the presence of primers 1 and 4 to produce a 1590 bp fragment, which was then digested with Pst I and 15 The CMV promoter fragment and the HIV-1 env Not I. fragment were then ligated together, and the entire subcloned into PPI4, which transcription unit eukaryotic shuttle vector that contains an ampicillin resistance gene, an SV40 origin of replication and a DHFR 20 gene whose transcription is driven by the ß-globin promoter. The final construct, PPI4-tPA-gp120LA, is shown in Figure 2.

The expression vector is then used as the prototype vector for the expression of gp120 proteins that are derived from other HIV-1 strains or mutated as described in the methods section. The vector was constructed so that unique Nar I and Not I sites flank the gp120 sequence, thus facilitating the removal of the gp120 gene cassette and the subsequent insertion of other gene cassettes (Figure 2).

- 2. Expression of HIV-1 gp120 in mammalian cells.
- a. <u>Transient expression</u>.

CosM5 cells grown in DMEM containing 10% fetal calf serum

were split to 75% confluence. On the following day, the cells were transfected for 16-20 hours with 10 micrograms of CsCl-purified PPI4-tPA-gp120<sub>LAI</sub> DNA by the standard CaPO<sub>4</sub> (5) precipitation technique. After transfection, fresh medium 5 was added to the cells. Analysis of the products synthesized 96-120 hours post-transfection was performed by radiolabelling the transfectants with 35S-cysteine for 12-18 hours, followed by precipitation of media using a CD4immunoglobulin-Protein A-Sepharose complex, followed by SDS-PAGE under reducing conditions (Figure 4).

## Stable expression.

Dhfr Chinese hamster ovary cells (CHO) were transfected with 20 micrograms of CsCl-purified DNA. Approximately 3-5 15 days post-transfection, cells were placed in selective medium (nucleoside-free alpha MEM containing 10% dialyzed fetal calf serum). Approximately 10-15 days post-selection, individual cell clones were picked. Media was analyzed for gp120 expression by radiolabelling the cells with 35Scysteine for 12-18 hours, followed by precipitation of media 20 using a CD4-immunoglobulin-Protein A-Sepharose complex. followed in turn by SDS-PAGE under reducing conditions The levels of gp120 in the media of these (Figure 6). clones were also quantitated (Figure 5) by ELISA performed The method involves coating 96-well plates 25 as follows. overnight with sheep polyclonal IgG against the highly conserved C-terminus of gp120 (D7234, Aalto Bioreagents). After washing, dilutions of a standard gp120 preparation in growth medium, or supernatant from the transfected cells, were incubated for 1 hour. 30 The plates were washed again, and incubated for one hour with a horseradish peroxidase-conjugated anti-gpl20 monoclonal Following a final wash, the antibody (9204, DuPont). peroxidase substrate OPD (DuPont) was added and the amount

of gp120 determined by comparing absorbance of unknowns with a standard curve. Standards were prepared from purified gp120 made in CHO cells, a small quantity of which was obtained from Celltech Ltd. Clones expressing the highest levels were subjected to successive rounds of amplification of the newly introduced DNA sequences in increasing concentrations of methotrexate. Stable CHO cell lines were thus generated which secrete at least 1 microgram/milliliter of HIV-1<sub>IAI</sub> gp120.

10

15

20

25

30

### 3. Construction of PPI4-tPA-qp120<sub>R-FL</sub>

a. The HIV-1<sub>LAI</sub> gp120 env nucleotide sequence in PPI4-tPA-gp120<sub>LAI</sub> was replaced by the nucleotide sequence encoding the mature gp120<sub>JR-FL</sub> protein. Using the polymerase chain reaction, the JR-FL sequences were amplified from pUC112-1 (27) using primer 5 (GATCGGCGCCAGAGTAGAAAAGTTGTGGGTCAC) and primer 4. The PCR fragment was digested with the restriction endonucleases Nar I and Not I, and the fragment subcloned in between the Nar I and Not I sites in PPI4-tPA-gp120<sub>LAI</sub> to generate PPI4-tPA-gp120<sub>R-FL</sub> (Figure 7).

### b. Transient expression.

CosM5 cells grown in DMEM containing 10% fetal calf serum were split to 75% confluence. On the following day, the cells were transfected for 16-20 hours with 10 micrograms of CsCl-purified PPI4-tPA-gp120<sub>JR-FL</sub> DNA by the standard CaPO<sub>4</sub> (5) precipitation technique. After transfection, fresh medium was added to the cells. Analysis of the products synthesized 96-120 hours post-transfection was performed by radiolabelling the transfectants with <sup>35</sup>S-cysteine for 12-18 hours, followed by precipitation of media using a CD4-immunoglobulin-Protein A-Sepharose complex, followed by SDS-PAGE under reducing conditions (Figure 4).

4. Construction of PPI4-tPA-gp120[A]-V3(1).

The V3 loop in tPA-gp120<sub>LAI</sub> consists of amino acids Cys<sub>306</sub> In the V3<sup>(-)</sup> mutant, the amino acids in through Cys, between these cysteines are replaced by the pentapeptide 5 sequence Thr-Gly-Ala-Gly-His. Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the V3 loop sequence in PPI4-tPA-gp120 $_{LAI}$  is altered using the mutagenic primer 6 (CTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAACATTAGTAGAGC) primer 7 (CTCGAGCATGCATTCGAAGCTCGCTGATC) as a selection Primer 7 changes a unique Xba I site in the 10 backbone of the parent PPI4 plasmid into a unique BstB I Briefly, the mutagenesis method requires incubating of the parent plasmid with the mutagenic primer and the selection primer, denaturing at 100°C for 3 minutes and then chilling on ice. In the presence of buffered deoxynucleo-15 tide triphosphates and T4 DNA polymerase, the primers are allowed to initiate the polymerization of one strand of plasmid DNA. T4 DNA ligase is used to seal the newly synthesized DNA strand to form a covalently closed circle. Hybrid plasmids are then transformed into a MutS strain of 20 E. coli that is deficient in mismatch repair. allowing for the growth of transformed cells, DNA is purified from the cells and digested with the selection restriction endonuclease, in this case Xba I. plasmids are cleaved by Xba I while the mutant plasmid 25 remains resistant to cleavage by virtue of the Xba I to BstB I conversion. Digested DNA is then used to transform E. coli, and colonies harboring the mutant plasmid are picked. Multiple mutagenic primers can be used in a single round of The amino acid sequence of the modified mutagenesis. 30 protein is shown in Figure 8.

5. Construction of PPI4-tPA-gp120 $_{\rm IR,FL}$ -V3 $^{(\cdot)}$ . The V3 loop in tPA-gp120 $_{\rm IR,FL}$  consists of amino acids Cys $_{293}$ 

through Cys<sub>377</sub>. In the V3<sup>(-)</sup> mutant, the amino acids in between these cysteines are replaced by the pentapeptide sequence Thr-Gly-Ala-Gly-His. Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the V3 loop sequence in PPI4-tPA-gp120<sub>JR-FL</sub> is altered using the mutagenic primer 6 (CTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAACATTAGTAGAGC) and primer 7 as a selection primer. The amino acid sequence of the modified protein is shown in Figure 9.

## 10 6. Construction of PPI4-tPA-gp120<sub>LAI</sub>-CD4<sup>(4)</sup>.

Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the selection primer 7. and the mutagenic primer 8 (CAATTTATAAACATGGTGCAGGAAGTAGG), Trp<sub>437</sub> of tpA-gp120<sub>LAI</sub>, which is in an equivalent position to the tryptophan residue in the HXBc2 strain of HIV-1, is mutated to a Val in the expression vector PPI4-tPA-gp120<sub>LAI</sub> to generate PPI4-tPA-gp120<sub>LAI</sub>-CD4<sup>(4)</sup>. The sequence for gp120<sub>LAI</sub>-CD4<sup>(4)</sup> is shown in Figure 12.

## 20 7. Construction of PPI4-tPA-gp120mm-CD4().

In a fashion similar to that described above, Trp<sub>424</sub> of tpA-gp120<sub>IR-FL</sub> is mutated to a Val in the expression vector PPI4-tPA-gp120<sub>IR-FL</sub> using the selection primer 7 and the mutagenic primer 9 (CAAATTATAAACATGGTGCAGGAAGTAGG) to generate PPI4-tPA-gp120<sub>IR-FL</sub>-CD4<sup>(-)</sup>. The sequence for gp120<sub>IR-FL</sub>-CD4<sup>(-)</sup> is shown in Figure 13.

## 8. Construction of PPI4-tPA-gp120, 1-V3()-CD4().

The tPA-gp120<sub>LAI</sub> double mutant, V3<sup>(4)</sup>-CD4<sup>(4)</sup>, is constructed by including the mutagenic primers 6 and 8, and the selection primer 7 simultaneously in the reaction tube with PPI4-tpA-gp120<sub>LAI</sub> as the DNA template. The final construct is named PPI4-tPA-gp120<sub>LAI</sub>-V3<sup>(4)</sup>-CD4<sup>(4)</sup>, and its sequence is shown in figure 10.

## 9. Construction of PPI4-tPA-qp120<sub>JR-FL</sub>-V3<sup>(4)</sup>-CD4<sup>(4)</sup>.

The tPA-gp120<sub>JR-FL</sub> double mutant, V3<sup>(·)</sup>-CD4<sup>(·)</sup>, is constructed by including the mutagenic primers 6 and 9, and the selection primer 7 simultaneously in the reaction tube with PPI4-tPA-gp120<sub>JR-FL</sub> as the DNA template. The final construct is named PPI4-tPA-gp120<sub>JR-FL</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup>, and its sequence is shown in figure 11.

## 10. Expression of mutant HIV-1 gp120 in mammalian cells.

## 10 a. <u>Transient expression</u>.

CosM5 cells grown in DMEM containing 10% fetal calf serum are split to 75% confluence. On the next day, the cells are transfected for 16-20 hours with 10 micrograms of CsClpurified mutant HIV-1 DNA by the standard CaPO, (5) precipitation technique. After transfection, fresh medium 15 is added to the cells. Analysis of the products synthesized 96-120 hours post-transfection is performed radiolabelling the transfectants with 35-cysteine for 12-18 hours, followed by precipitation of media using a sheep polyclonal IgG against the highly conserved C-terminus of 20 gp120.

### b. Stable expression.

Dhfr Chinese hamster ovary cells (CHO) are transfected with 25 20 micrograms of CsCl-purified DNA encoding the native or mutant HIV-1 gp120 glycoproteins. Approximately 3-5 days post-transfection, cells are placed in selective medium (nucleoside-free alpha MEM containing 10% dialyzed fetal Approximately 10-15 days post-selection, calf serum). individual cell clones are picked. Media is analyzed for 30 gp120 expression by radiolabelling the cells with 35Scysteine for 12-18 hours, followed by quantitative immunoprecipitation of media using a sheep polyclonal IgG against the highly conserved C-terminus of gp120, followed

SDS-PAGE under reducing conditions. inby Alternatively, one can quantitate the level of gp120 by ELISA performed as follows. The method involves coating 96well plates overnight with sheep polyclonal IgG against the 5 highly conserved C-terminus of qp120 (D7234, Bioreagents). After washing, dilutions of a standard gp120 preparation in cell growth medium, or supernatant from the stably-transfected cells, are incubated for 1 hour. plates are washed again, and incubated for one hour with a human MoAb (F105, AIDS Research & Reference Reagent Program, 10 No. 857). The plates are washed again, and incubated again for 1 hour with a horseradish-peroxidase-conjugated goat anti-human IgG (Cappel). Following a final wash, the peroxidase substrate OPD (DuPont) is added and the amount of gp120 determined by comparing absorbance of unknowns with a 15 standard curve. Standards are prepared from purified qp120 made in CHO cells, a small quantity of which is obtained from Celltech Ltd. Clones expressing the highest levels are subjected to successive rounds of amplification of the newly introduced DNA sequences in increasing concentrations of 20 Stable CHO cell lines are thus generated methotrexate. which secrete at least 1 microgram/milliliter of mutant HIV-1 gp120.

## 25 11. Purification of HIV-1 gp120 proteins.

A one-step immunoaffinity procedure is used to purify the recombinant gp120 molecules described. Briefly, culture supernatant is collected and clarified by centrifugation. An immunoaffinity column consisting of a matrix coupled to a sheep polyclonal anti-gp120 IgG (D7234, Aalto Bioreagents) directed against the highly conserved C-terminal end (APTKAKRRVVQREKR) of gp120 is used to specifically adsorb gp120 from the cell culture media. This antisera recognizes native gp120, the V3 loop deletion mutants, and the CD4<sup>(-)</sup>

WO 94/22477 PCT/US94/03282

mutants since the C-terminal ends of these molecules remain unaltered. The bound gp120 is then eluted with 2M MgCl<sub>2</sub>, concentrated by Amicon filtration, and dialyzed into 10 mM HEPES, pH 7.0. The purity of the proteins is determined by SDS-PAGE and silver staining.

43

Characterization of recombinant HIV-1 gp120 proteins. The purified glycoproteins are subjected to extensive biochemical and immunologic characterization. The integrity of the proteins is monitored by SDS-PAGE and silver staining 10 under reducing and non-reducing conditions. glycoproteins are deglycosylated by treatment with the enzyme N-glycosidase F which cleaves N-linked oligosaccharides, and are assayed by SDS-PAGE and silver staining to monitor molecular weight shifts. 15 The glycoproteins are also tested for reactivity with several well characterized anti-gp120 monoclonal antibodies that recognize both linear and discontinuous epitopes. binding affinity to sCD4 is estimated using an ELISA assay.

The purified proteins HIV-1 gp120<sub>LAI</sub>, gp120<sub>LAI</sub>-V3<sup>(\*)</sup>, gp120<sub>LAI</sub>-V3<sup>(\*)</sup>

)-CD4<sup>(\*)</sup>, gp120<sub>IR-FL</sub>, gp120<sub>IR-FL</sub>-V3<sup>(\*)</sup>, and gp120<sub>IR-FL</sub>-V3<sup>(\*)</sup>-CD4<sup>(\*)</sup>, were tested for their ability to bind cell surface human CD4. DG44 #3 cells, a recombinant cell line designed to express human CD4 on the membrane surface, were grown in T flasks and trypsinized. 5 X 10<sup>5</sup> cells/experiment were aliquoted into FACS buffer (PBS + 2\* BSA and 0.1\* NaN<sub>3</sub>), washed several times in the same buffer, and then incubated with 100 ul of a solution of purified gp120 protein at 5ug/ml in FACS buffer at 37°C for 2 hr. The cells were washed in FACS buffer, and then incubated in 100 ul solution containing 5ug/ml sheep polyclonal IgG against the highly conserved C-terminus of gp120 in FACS buffer at 37°C for 2 hr. The cells were washed in FACS buffer then incubated in 100 ul

20

15

20

25

solution containing FITC-labeled rabbit anti-sheep IgG polyclonal antibody at 37°C for 2 hr. The cells were washed with FACS buffer and then resuspended in 500 ul FACS buffer. The cells were then analyzed on a Becton Dickinson FACScan according to the manufacturer's instructions. As a control for expression of CD4 on the DG44 #3 cells, FITC-labeled OKT4A (Becton Dickinson) was used.

# 13. A protocol for inoculation of animals with the mutant HIV-1 gp120 envelope qlycoproteins.

Alum is used as an adjuvant during the inoculation series. The inoculum is prepared by dissolving the mutant HIV-1 gp120 envelope glycoprotein antigen in physiologic saline at a final antigen concentration of 100 ug/ml. Preformed alum (aluminum hydroxide gel) is added to the solution to a final level of 500 ug/ml aluminum. The antigen is allowed to adsorb onto the alum gel for two hours at room temperature. Following adsorption, the gel with the antigen is washed twice with physiologic saline and resuspended in the saline to a protein concentration of 100 ug/ml.

Monkeys and/or Guinea Pigs are individually inoculated with four 100 ug doses of the mutant HIV-1 gp120 envelope glycoprotein antigen adsorbed onto alum. Each dose is injected intramuscularly. The doses are delivered one or five months apart (week 0, 4, 8 and 28). the animals are bled at intervals of two or four weeks. Serum samples are prepared from each bleed to assay for the development of specific antibodies as described in the subsequent sections.

30

35

# 14. Analysis of sera for anti-mutant HIV-1 gp120 envelope glycoprotein IqG antibodies.

Each serum sample is analyzed by ELISA. Polystyrene microtiter plates are coated with 0.5 ug per well of pure mutant HIV-1 gp120 envelope glycoprotein in phosphate-

15

20

25

buffered physiological saline (PBS) at 4°C. Each well is then washed with PBS containing 0.5% TWEEN-20 (PBS-TW). Test serum, diluted serially in PBS-TW, is added to the mutant HIV-1 gp120 envelope glycoprotein-containing wells 5 and allowed to react with the adsorbed mutant HIV-1 gp120 envelope glycoprotein for one hour at 37°C. The wells are then washed extensively in PBS-TW. Each well then receives 0.1% p-nitrophenyl phosphate in 10% diethanolamine, pH 9.8, containing 0.5 mM MgCl<sub>2</sub>.6H<sub>2</sub>0. The ensuing reaction is allowed to proceed at room temperature for 30 minutes, at which time it is terminated by the addition of 3.0 N NaOH.

The greater the interaction of antibodies in the test serum with the mutant HIV-1 gp120 envelope glycoprotein, the greater is the amount of alkaline phosphatase bound onto the well. The phosphatase enzyme mediates the breakdown of pnitrophenyl phosphate into a molecular substance which absorbs light at a wavelength of 405 nm. Hence, there exists a direct relationship between the absorbance at 405 nm of light at the end of the ELISA reaction and the amount of mutant HIV-1 gp120 envelope glycoprotein-bound antibody. All animals inoculated with mutant HIV-1 gp120 envelope glycoprotein whose serum reacts specifically with the mutant HIV-1 gp120 envelope glycoprotein in the ELISA have a positive antibody response against mutant HIV-1 gp120 envelope glycoprotein.

## Analysis of sera for activity which specifically neutralizes HIV-1 infectivity.

Virus-neutralizing activity is determined with an assay based on the use of multiplicity curves in which the ratio of infectious virus surviving antibody treatment  $(V_n)$  is compared to infectious virus in uninhibited cultures (Va) at various dilutions of antisera. The neutralization titer of

the sera is then interpolated as that sera dilution which yields one log reduction in infectious titer (i.e.,  $V_{\rm n}/V_{\rm o}$  = Briefly, 4-fold dilutions of virus (laboratoryadapted and primary isolates) are prepared to yield 5 infectious doses of 0.1 to 100 TCID, (Tissue Culture Infection Dose) in 20 ul. Serial 3-fold dilutions of sera are also prepared and 20 ul of each serum dilution are incubated with each dilution of virus in duplicate for 60 minutes at room temperature in a 96-well microtiter plate. 10 20 ul of AA5 cells (PHA stimulated PBMCs for primary HIV-1 isolates) are then added to the serum/virus mixtures. Cells are cultured for 7 days by the addition of fresh medium every other day. On the seventh day, supernatant from each well is removed and tested for the presence of reverse transcriptase (RT). Infection in each well is then scored 15 as either positive or negative based on the RT counts, and the infectious dose of virus in each treatment group is calculated using the Reed and Muench (28) formula. The neutralization titers represent the reciprocal dilution required to reduced infectious dose of virus by one log. The above culture time is for the prototypic HIV-1, AI isolate tested on the AA5 cell line. In the case of primary isolates, the termination date is usually 11-14 days. Culture conditions for PBMCs is not as demanding since doubling time is restricted. In the case of PBMCs, one day PHA stimulations are used at a final concentration of 1.5 X 106/ml on day 0. Half that number of fresh PBMCs are then added again on days 4 and 8. This multiple addition of PBMCs is meant to amplify virus output upon successful infection so that the readout RT signal is strong. 30 the final readout titer for the primary isolate/PBMC is the reciprocal serum dilution which reduces infectious titer by one log.

## 16. Passive hyperimmune therapy.

Non-HIV-1-infected humans are immunized with the mutant HIV1 gp120 envelope glycoprotein antigens according to a
protocol similar to that described above in section 12. For
5 passive hyperimmune therapy in HIV-1-infected individuals,
blood plasma is taken from mutant HIV-1 gp120 envelope
glycoprotein immunized, non-HIV-1-infected human donors
whose plasma has high levels of neutralizing antibodies.
The plasma is pooled from several donors, purified to remove
nonimmunoglobulin proteins and is then sterilized to kill
any other viruses or pathogens. The treated plasma is then
injected into individuals infected with HIV-1, with repeated
injections every week, every two weeks, or every month.

### Results

Eukaryotic expression vectors designed to express high levels of HIV-11AI gp120 and HIV-11R-SL gp120 were constructed. The CMV MIE promoter/enhancer was used to drive the transcription of a gene fusion consisting of the human tPA signal sequence fused to mature gp120 (Figures 2 and 7). The complete sequence of the transcription unit from the Hinc II site of the CMV promoter/enhancer to the Not I site just 3' from the stop codon in gp120 is shown in figure 3. 10 This vector was used to transfect COSM5 cells in a transient assay. The transfected cells were labeled with 35-cysteins and the media immunoprecipitated with a CD4-immunoglobulin-Protein A-Sepharose complex. The precipitated products were analyzed using a reducing 10% 15 SDS-PAGE autoradiography (Figure 4). A 120 kD band was detected when PPI4-tPA-gp120LAI was used to transfect COS cells (lane 3). A band migrating with a slightly lower molecular mass was detected when PPI4-tPA-gp120 $_{\mbox{\scriptsize IR-FL}}$  was used to transfect COS cells (lane 4). No radiolabeled products were detected in 20 the mock infected cells. Using a sheep polyclonal antibody directed against the highly conserved C-terminal end of HIV-1 gp120 in an ELISA assay, the level of expression of HIV-1 qp120 was determined to be 2350 ng/ml.

25

30

The PPI4-tPA-gp120<sub>LA</sub> vector was then used to stably transfect the dhfr CHO cell line DXB11. Two days post-transfection, the cells were plated at low density in nucleoside-free medium. Eight days post-transfection, surviving clones were isolated and expanded. Individual primary transfectants were tested for gp120 expression using the ELISA method described in the methods section. Several primary CHO transfectants expressed significant quantities (10-120 ng/ml) of gp120 (Figure 5). Three of the highest

expressing clones were then subjected to increasing concentrations of methotrexate in order to amplify, in tandem, the copy number of the dhfr and gp120 genes. Cell lines were established that express high levels of gp120 with rates of secretion greater than 1 mg/liter. These were then used to purify gp120 to homogeneity.

Six CHO cell lines were established, using the procedures described in the methods sections, that express high levels of the following proteins: HIV-1 gp120<sub>LAI</sub>, gp120<sub>LAI</sub>-V3<sup>(·)</sup>, gp120<sub>LAI</sub>-V3<sup>(·)</sup>, gp120<sub>JR-FL</sub>, gp120<sub>JR-FL</sub>-V3<sup>(·)</sup>, and gp120<sub>JR-FL</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup>. Metabolic labeling of these cells with <sup>35</sup>S-cysteine followed by immunoprecipitation with the human monoclonal antibody F105 and analyzed by SDS-PAGE and autoradiography showed the presence of the gp120 proteins in the culture supernatant (Figure 14). From these cell lines the gp120 proteins were purified to homogeneity. Analysis by SDS-PAGE followed by silver-staining showed the purity of these proteins to be greater than 90% (Figure 15).

20

25

10

It was shown by FACScan analysis that the two CD4 binding mutants  $HIV-1gp120_{IAI}-V3^{(\cdot)}-CD4^{(\cdot)}$  and  $HIV-1gp120_{IR-FL}-V3^{(\cdot)}-CD4^{(\cdot)}$  had no appreciable binding to recombinant cell lines designed to express high levels of human CD4 on their membrane surface (Figure 16, panel 4 and data not shown, respectively).

### Discussion

The advantage of using the mutant HIV-1 gp120 envelope glycoproteins as immunogens is that these proteins will not elicit an immune response against the V3 loop, a highly immunodominant epitope on gp120. This is significant because the V3 loop may skew the humoral immune response away from discontinuous epitopes in the CD4-binding site. Mutant HIV-1 gp120 envelope glycoproteins having partial and total v3 loop deletions have been made (30). Deletion of the V3 loop therefore exposes the CD4-binding site to the immune system. allowing the immune system to mount a response against this critical region (18). Another advantage of using the mutant HIV-1 gp120 envelope glycoprotein as an immunogen is that it 15 has significantly reduced affinity for cell surface CD4. An efficient humoral immune response depends on the binding of antigen to B cell surface immunoglobulin. The presence of the high-affinity CD4 receptor on large numbers of cells in the body may significantly diminish the ability of native 20 qp120 to induce an effective humoral immune response. rationale of mutating gp120 at the CD4 binding site is to redirect the mutant HIV-1 gp120 envelope glycoprotein away from cell surface CD4 toward immunoglobulin-bearing B cells. thereby allowing the immune system to mount a response 25 against, inter alia, the CD4-binding site.

### References

1. Klatzmann, D.R., et. al. (1990) Immunodeficiency Reviews 2, 43-66.

5

- Lasky, L.A., et. al. (1987) Cell 50, 975-985.
- 3. Maddon, P.J., et. al. (1986) Cell 47, 333-348.
- 10 4. Maddon, P.J., et. al. (1988) Cell 54, 865-874.
  - 5. Maddon, P.J., et. al. (1985) Cell 42, 93-104.
- 6. Maddon, P.J., et. al. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159.
  - 7. Richardson, N.E., et. al. (1988) Proc. Natl. Acad. Sci. U.S.A. 85, 6102-6106.
- 20 8. Chao, B.H., et. al. (1989) J. Biol. Chem. 264, 5812-5817.
  - 9. Arthos, J., et. al. (1989) Cell 57, 469-481.
- 25 10. Wang, J., et. al. (1990) Nature 348, 411-418.
  - 11. Ryu, S.-E., et. al. (1990) Nature 348, 419-426.
- 12. Leonard, C.K., et. al. (1990) J. Biol. Chem. 265, 10373-10382.
  - 13. Earl, P.L., et. al. (1990) Proc. Natl. Acad. Sci.
    U.S.A. 87, 648-652.
- 35 14. Helseth, E., et. al. (1991) J. Virol. 65, 2119-2123.

WO 94/22477 PCT/US94/03282

52

- 15. Bolognesi, D.P. (1990) TIBTech 8, 40-45.
- 16. Olshevsky, U., et. al. (1990) J. Virol. 64, 5701-5707.
- 5 17. Steimer, K.S., et. al. (1991) AIDS 5, S135-143.
  - 18. Wyatt, R., et. al. (1992) J. Virol. 66, 6997-7004.
- Zolla-Pazner, S., et. al. (1992) Sem. in Virology 3,
   203-211.
  - 20. Steimer, K.S., et. al. (1991) Science 254, 105-108.
  - 21. Pollard, S.R., et. al. (1992) EMBO J. 11, 585-591.
- 22. Okayama, H. (1983) Mol. Cell. Biol. 3, 280-289.

15

20

- 23. Maniatis, T., et. al. (1990) Molecular Cloning, Vol. 1-3.
- 24. Thomsen, D.R., et. al. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 659-663.
- 25. Pennica, D., et. al. (1983) Nature 301, 214-221.
  - 26. Wain-Hobson, S., et. al. (1985) Cell 40, 9-17.
    - 27. Koyanagi, Y. (1987) Science 236, 819-822.
- 30 28. Reed, L.J. (1938) Am. J. Hyg., 27, 493-497.
  - 29. Cohn, E.J. et al., (1944) J. Clin. Invest. 23, 417-432.
- 30. Shiow-Her, C., et al. (1992) J. of Cellular Biochem., 35 Supplement 16E, Abstrtact Q105.

#### SEQUENCE LISTING

- (1) GENERAL INFORMATION: (i) APPLICANT: Progenics Pharmaceuticals, Inc. (ii) TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES (iii) NUMBER OF SEQUENCES: 29 (iv) CORRESPONDENCE ADDRESS: (A) ADDRESSEE: Cooper & Durham (B) STREET: 30 Rockefeller Plaza (C) CITY: New York (D) STATE: New York (E) COUNTRY: USA (F) ZIP: 10112 (v) COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: Patentin Release #1.24 (vi) CURRENT APPLICATION DATA: (A) APPLICATION NUMBER: (B) FILING DATE: (C) CLASSIFICATION: (vii) PRIOR APPLICATION DATA: (A) APPLICATION NUMBER: US 08/037,816 (B) FILING DATE: 26-MAR-1993 (viii) ATTORNEY/AGENT INFORMATION: (A) NAME: White, John P. (B) REGISTRATION NUMBER: 28,678 (C) REFERENCE/DOCKET NUMBER: 41190-A-PCT/JPW/AJM (ix) TELECOMMUNICATION INFORMATION: (A) TELEPHONE: (212) 977-9550 (B) TELEFAX: (212) 664-0525 (C) TELEX: 422523 COOPUI (2) INFORMATION FOR SEQ ID NO:1: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 45 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: Linear (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
  - You Xam Xam Cys Xam Ile Xam Xam Xam Xam Xam Xam Trp Xam Xam Xam 1 10 15

  - Ser xaa xaa Thr Gly xaa xaa xaa Arg xaa Gly xaa

- (2) INFORMATION FOR SEQ ID NO:2:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 45 amino acids (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Thr Leu Pro Cys Arg Ile Lys Gln Phe Ile Asn Met Trp Gln Glu Val 1 5 10 15

Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser 20 25 30

Ser Asn lie Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly 35 40 45

- (2) INFORMATION FOR SEQ ID NO:3:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 45 amino acids
    - (B) TYPE: amino acid -
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Net Trp Gln Glu Val

Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser 20 25 30

Ser Asn Ile Thr Gly Leu teu Leu Thr Arg Asp Gly Gly 35 40 45

- (2) INFORMATION FOR SEQ ID NO:4:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 45 base pairs
    - (8) TYPE: nucleic acid
    - (C) STRAMDEDNESS: single (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GATCCTGCAG TCACCGTCCT TGACACGATG GATGCAATGA AGAGA

. 45

- (2) INFORMATION FOR SEQ ID NO:5:
  - (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 36 base pairs
    - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
AAGTCTTCTC CTCGGTCTTG TCTTTTTAAC ACCCAG	36
(2) INFORMATION FOR SEQ ID NO:6:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 36 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: DNA (genomic)	
( ), metalogic ( ) and (gendance)	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
TTCAGAAGAG GAGCCAGAAC AGAAAAATTG TGGGTC	36
(2) INFORMATION FOR SEG ID NO:7:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 39 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:	
GGAAAAAGC GGCCGCTCAT TTTTCTCTCT GCACCACTC	39
(2) INFORMATION FOR SEQ ID NO:8:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 33 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	., .
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	
CATCGGCGCC AGAGTAGAAA AGTTGTGGGT CAC	33
(2) INFORMATION FOR SEQ ID NO:9:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 49 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	

(ii) MOLECULE TYPE: DNA (genomic)

56 .

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:	
CTGTAGAAAT TAATTGTACA GGTGCTGGAC ATTGTAACAT TAGTAGAGC	49
(2) INFORMATION FOR SEQ ID NO:10:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 29 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:	
CTCGAGCATG CATTCGAAGC TCGCTGATC	29
(2) INFORMATION FOR SEQ ID NO:11:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 29 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:	
CAATTTATAA ACATGGTGCA GGAAGTAGG	29
(2) INFORMATION FOR SEG ID NO:12:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 29 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:	
CAAATTATAA ACATGGTGCA GGAAGTAGG	29
2) INFORMATION FOR SEQ ID NO:13:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 3125 base pairs  (B) TYPE: nucleic acid  (C) STRAMDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 15553115 (D) OTHER INFORMATION:	

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TIGACATIGA TTATTGACTA GITATTAATA GTAATCAATT ACGGGGTCAT TAGTTCATAG	60
CCCATATATG GAGTTCCGCG TTACATAACT TACGGTAAAT GGCCCGCCTG GCTGACCGCC	120
CAACGACCCC CGCCCATTGA CGTCAATAAT GACGTATGTT CCCATAGTAA CGCCAATAGG	180
GACTITICCAT IGACGICAAT GGGIGGACTA TITACGGIAA ACTGCCCACT IGGCAGTACA	240
TCAAGTGTAT CATATGECAA GTACGCCCCC TATTGACGTC AATGACGGTA AATGGCCCGC	300
CTGGCATTAT GCCCAGTACA TGACCTTATG GGACTTTCCT ACTTGGCAGT ACATCTACGT	360
ATTAGTCATC GCTATTACCA TGGTGATGCG GTTTTGGCAG TACATCAATG GGCGTGGATA	420
GCGGTTTGAC TCACGGGGAT TTCCAAGTCT CCACCCCATT GACGTCAATG GGAGTTTGTT	480
TTGGCACCAA AATCAACGGG ACTTTCCAAA ATGTCGTAAC AACTCCGCCC CATTGACGCA	540
AATGGGCGGT AGGCGTGTAC GGTGGGAGGT CTATATAAGC AGAGCTCGTT TAGTGAACCG	600
TCAGATCGCC TGGAGACGCC ATCCACGCTG TTTTGACCTC CATAGAAGAC ACCGGGACCG	660
ATCCAGCCTC CGCGGCCGGG AACGGTGCAT TGGAACGCGG ATTCCCCGTG CCAAGAGTGA	720
CGTAAGTACC GCCTATAGAC TCTATAGGCA CACCCCTTTG GCTCTTATGC ATGCTATACT	780
GTTTTTGGCT TGGGCCAACA CCCCGTCCTA GATAGGTGAT GGTATAGCTT AGCCTATAGG	840
TGTGGGTTAT TGACCATTAT TGACCACTCC CCTATTGGTG ACGATACTTT CCATTACTAA	900
TCCATAACAT GGCCGCTCTT TGCCACAACT ATCTCTATTG GCTATATGCC AATACTCTGT	960
CCTTCAGAGA CTGACACGGA CTCTGTATTT TTACAGGATG GGGTCCCATT TATTATTTAC	1020
AMATTCACAT ATACAACAAC GCCGTCCCCC GTGCCCGCAG TTTTTATTAA CATGCGGGAT	1080
CTCCACGCGA ATCTCGGGTA CGTGTTCCGG ACATGGGCTC TTCTCCGGTA GCGGCGGAGC	1140
TECACATECG AGEETGTEEE ATGECEATGE CTECAGEGGE TEATGGTEGE TEGGEAGETE	1200
CTTGCTCCTA ACAGTGGAGG CCAGACTTAG GCACAGGACA ATGCCCACCA CCACCAGTGT	1260
GCCGCACAAG GCCGTGGCGG TAGGGTATGT GTCTGAAAAT GAGCTCGGAG ATTGGGCTCG	1320
CACCGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG GCAGCTGAGT	1380
TGTTGTATTC TGTAGAGTTG GAGGTAACTC CCGTTGCGGT GCTGTTAACG GTGGAGGGCA	1440
GTGTAGTCTG AGCAGTACTE GTTGCTGCCG EGCGCGCCAC CAGACATAAT AGCTGACAGA	1500
CTANCAGACT GTTCCTTTCC ATGGGTCTTT TCTGCAGTCA CCGTCCTTGA CACG ATG Met 1	1557
GAT GCA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG CTG TGT GGA GCA ASP Ala Het Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly Ala 5 10 15	1605
GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA GGC Val Phe Val Ser Pro Ser Gin Glu Ile His Ala Arg Phe Arg Arg Gly 20 25 30	1653
GCC AGA ACA GAA AAA TTG TGG GTC ACA GTC TAT TAT GGG GTA CCT GTG Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val 35 40 45	1701
TGG AAG GAA GCA ACC ACC ACT CTA TTT TGT GCA TCA GAT GCT AAA GCA	1749

Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala 50 55 60 65	
TAT GAT ACA GAG GTA CAT AAT GTT TGG GCC ACA CAT GCC TGT GTA CCC Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro 70 75 80	179
ACA GAC CCC AAC CCA CAA GAA GTA GTA TTG GTA AAT GTG ACA GAA AAT Thr Asp Pro Asn Pro Gin Giu Val Val Leu Val Asn Val Thr Giu Asn 85 90 95	184
TTT AAC ATG TGG AAA AAT GAC ATG GTA GAA CAG ATG CAT GAG GAT ATA Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile 100 105 110	1893
ATC AGT TTA TGG GAT CAA AGC CTA AAG CCA TGT GTA AAA TTA ACC CCA lie Ser Leu Trp Asp Gin Ser Leu Lys Pro Cys Val Lys Leu Thr Pro 115 120 125	1941
CTC TGT GTT AGT TTA AAG TGC ACT GAT TTG GGG AAT GCT ACT AAT ACC Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Ash Ala Thr Ash Thr 135 140 145	1989
AAT AGT AGT AAT ACC AAT AGT AGT AGC GGG GAA ATG ATG GAG AAA Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu Lys 150 155 160	2037
GGA GAG ATA AAA AAC IGC ICT ITC AAT ATC AGC ACA AGC ATA AGA GGT Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg Gly 165 170 175	2085
AAG GTG CAG AAA GAA TAT GCA TTT TTT TAT AAA CTT GAT ATA ATA CCA Lys Val Gin Lys Giu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile Pro 180 185 190	2133
ATA GAT AAT GAT ACT ACC AGC TAT ACG TTG ACA AGT TGT AAC ACC TCA  Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr Ser  195 200 205	2181
GTC ATT ACA CAG GCC TGT CCA AAG GTA TCC TTT GAG CCA ATT CCC ATA Val lie Thr Gin Ala Cys Pro Lys Val Ser Phe Glu Pro lie Pro lie 210 215 220. 225	2229
CAT TAT TGT GCC CCG GCT GGT TTT GCG ATT CTA AAA TGT AAT AAG His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn Lys 230 235 240	2277
ACG TTC AAT GGA ACA GGA CCA TGT ACA AAT GTC AGC ACA GTA CAA TGT Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys 245 250 255	2325
ACA CAT GGA ATT AGG CCA GTA GTA TCA ACT CAA CTG CTG TTG AAT GGC Thr His Gly Ile Arg Pro Val Vel Ser Thr Gln Leu Leu Leu Asn Gly 260 265 270	2373
AGT CTA GCA GAA GAA GAG GTA GTA ATT AGA TCT GCC AAT TTC ACA GAC Ser Leu Ala Glu Glu Val Val Ile Arg Ser Ala Asn Phe Thr Asp 275 280 285	2421
AAT GCT AAA ACC ATA ATA GTA CAG CTG AAC CAA TCT GTA GAA ATT AAT Asn Ala Lys Thr Ite Ite Val Gln Leu Asn Gln Ser Val Glu Ite Asn 290 295 300 305	2469
TGT ACA AGA CCC AAC AAC AAT ACA AGA AAA AGT ATC CGT ATC CAG AGG Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gin Arg 310 315	2517
GGA CCA GGG AGA GCA TTT GTT ACA ATA GGA AAA ATA GGA AAT ATG AGA Gly Pro Gly Arg Ala Phe Val Thr Ile Gly Lys Ile Gly Asn Met Arg	2565

			325	5				330	)				335			
			Cys			r AGT e Ser		Ala					Thr			2613
		ALE				AGA Arg 360	Glu									2661
A10 Ile 370	Phe	AAC Lys	CAA Gln	TCC Ser	TCA Ser 375	GGA	GGG	GAC	CCA Pro	GAA Glu 380	ATT	GTA Val	ACG Thr	CAC	AGT Ser 385	2709
TTT Phe	AAT Asn	TGT Cys	GGA	6GG 6ly 390	Glu	TTT Phe	TTC Phe	TAC	TGT Cys 395	AAT Asn	TCA Ser	ACA Thr	CAA Gln	CTG Leu 400	TTT Phe	2 <i>1</i> 57
						AGT Ser			Ser							2805
			Ser			ATC										2853
						GTA Val 440										2901
						TCA Ser										2949
Arg	Asp	Gly	Gly	АSП 470	Asn	AAC Asn	Asn	Gly	Ser 475	Glu	Ite	Phe	Arg	Pro 480	Gly	2997
Gly	Gly	Asp	Met 485	Arg	Asp	AAT Asn	Trp	Arg 490	Ser	Glu	Leu	Tyr	Lys 495	Tyr	Lys	3045
Val	Val	Lys 500	lle	Glu	Pro		Gly 505	Val	Ala	Pro						3093
				AGA Arg	Glu	AAA Lys 520	T GA	GCGG	CCGC	:						3125

- (2) INFORMATION FOR SEQ ID NO:14:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 520 amino acids
    - (B) TYPE: amino acid
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly

Ala Vai Phe Vai Ser Pro Ser Gin Giu 11e His Ala Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro

49

Val Trp Lys Glu Ala Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 55 70 75 80

Pro Thr Asp Pro Asn Pro Gin Giu Val Val Leu Val Asn Val Thr Giu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Met Glu 145 150 155 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr. Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro 210 215 220

Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gin Leu Leu Leu Asn 260 265 270

Gly Ser Leu Ala Glu Glu Glu Val Val Ile Arg Ser Ala Asn Phe Thr 275 280 285

Asp Asn Ala Lys Thr Ile Ile Val Gin Leu Asn Gin Ser Val Glu Ile 290 295 300

Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gln 305 310 320

Arg Gly Pro Gly Arg Ala Phe Val Thr Ile Gly Lys Ile Gly Asn Met 325 335

Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Ala Thr Leu 340 345 350

Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln Phe Gly Asn Asn Lys Thr 355 360 365

lie lie Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu lie Val Thr His 370 380

Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu 385 390 395 400

Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn

	61					
	405	410	415			
Asn Thr Glu Gl 42		ile Thr Leu Pr 425	o Cys Arg ile Lys Gln 430			
Phe Ile Asn Me 435	t Trp Gin Giu	Val Gly Lys Al 440	a Met Tyr Ala Pro Pro 445			
Ite Ser Gly Gl	n lle Arg Cys 455	Ser Ser Asn Il	e Thr Gly Leu Leu Leu 460			
Thr Arg Asp Gly	y Gly Asn Asn 470	Asn Asn Gly Ser 47	Glu Ile Phe Arg Pro 480			
Gly Gly Gly Asp	Met Arg Asp 485	Asn Trp Arg Ser 490	Glu Leu Tyr Lys Tyr 495			
Lys Val Val Lys 500		Leu Gly Val Ala 505	Pro Thr Lys Ala Lys 510			
Arg Arg Val Val 515		Lys 520				
(2) INFORMATION	FOR SEQ ID N	0:15:				
(A) L (B) T (C) S	CE CHARACTERI ENGTH: 1532 b YPE: nucleic TRANDEDNESS:: OPOLOGY: line	ase pairs acid single				
(ii) MOLECUI	LE TYPE: DNA	(genomic)				
(B) L	E: AME/KEY: CDS DCATION: 115 THER INFORMATI					

(xi)	SEQUENCE	DESCRIPTION:	SEQ	ID	NO:15:
------	----------	--------------	-----	----	--------

					GTG Val			48
					CAT His			96
					GTC Val			144
					TGT Cys			192
					GCC Ala 75			240
					TTG Leu			288
					GAA Glu			336

				Le				ln S			NG CC /s Pr			Lys			384
	.o r		_				eu A:				IT GT		n Ala			ACC Thr	432
	r A						y Tr				A GG/ g Gl: 15!	y Glu					480
						r In					T GAO P Gio D						528
					LY					l Pro	A ATA						576
		r							o The		A GTC - Val						624
Pro	21	s 1	le	Ser	Phe	e Glu	21:	o 116	Pro	ı ile	CAT His	7yr 220	Cys	Ala	Pro	Ala	672
Gly 225	Ph	e A	la	He	Leu	230	Cy:	s Asr	) Asp	Lys	thr 235	Phe	Asn	Gly	Lys	Gly 240	720
Pro	Cy	5 L	ys	Asn	Val 245	Ser	Thr	· Val	Gin	250		His	Gly	Ile	Arg 255	Pro	768
Val	Va	l S	er :	Thr 260	Gln	Leu	Leu	Leu	265	Gly	AGT Ser	Leu	Ala	Glu 270	Glu	Glu	816
/al	Val	2 <sup>-</sup>	le / 75	Arg	Ser	Asp	ASN	280	Thr	Asn	AAT	Ala	Lys 285	Thr	Ile	He	864
/a l	Glr 290	) )	eu l	Lys	Glu	Ser	Val 295	Ģļu	He	Asn	Cys	Thr 300	Arg	Pro	Asn	Asn	912
15n 105	Thr	· Aı	rg L	Lys	Ser	1 l e 310	His	Tle	Gly	Pro	GGG Gly 315 GCA	Arg	Ala	Phe	Tyr	Thr 320	960
'nr	Gly	G	lu I	ile :	11e 325	Gly	Asp	Ile	Arg	Gln 330	Ala	His	Cys	Asn	11e 335	Ser	1008
rg	Ala	Ly	rs T	140	Asn	Asp	Thr	Leu	Lys 345	Gin	ATA	Val	Ile	L <i>ys</i> 350	Leu	Arg	1056
ŧυ	Gln	Ph 35	e G 5	ilu /	Asn	Lys	Thr	11e 360	Val	Phe	AAT ASN TGT	His	Ser 365	Ser	Gly	Gly	1104
sp	Pro <b>370</b>	Gl	u l	le \	/a l	Met	His 375	Ser	Phe	Asn	Cys	Gly 380	Gly	Glu	Phe		1152

1 y 38	r Cy 5	s As	n Se	r Th	r Glr 390	) )	) Pho	e Asr	sei	7 Thi		) Asn	ASF	Asn	1hr 400	
GA	A GG	G TC y Se	A AA r Asi	T AA( n Asi 405	C ACT	GAA	GGA Gly	AAT ASN	ACT Thr 410	· lie	ACA Thr	CTC	CCA Pro	TGC Cys 415	AGA Arg	1248
ATA	A AA	G CA	A ATT	e Ile	AAC Asn	ATG Met	TGG Trp	i i AG Gln 425	GAA	GTA Val	GGA Gly	AAA Lys	GCA Ala 430	ATG Met	TAT Tyr	1296
			lie		GGA Gly											1344
		Leu			GAT Asp											1392
11C Phe 465	AGA Arg	CCT Pro	GGA Gly	GGA Gly	GGA Gly 470	GAT ASP	ATG Met	AGG Arg	GAC ASP	AAT Asn 475	TGG Trp	AGA Arg	AGT Ser	GAA Glu	TTA Leu 480	1440
TAT Tyr	AAA Lys	TAT Tyr	AAA Lys	GTA Val 485	GTA Val	A#A Lys	ATT	Glu	CCA Pro 490	TTA Leu	GGA Gly	GTA Val	Ala	266 Pro 4 <b>9</b> 5	ACC Thr	:488
					GTG (		Gln				T GA	GCGG	CCGC			1532

### (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 507 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Ash Pro Gln Glu Val Val Leu Glu Ash Val Thr Glu 85 90 95

His Phe Asn Het Trp Lys Asn Asn Met Val Glu Gln Het Gln Glu Asp 100 \_ 105 110

lie lie Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Thr Leu Asn Cys Lys Asp Val Asn Ala Thr Asn Thr 130 135 140

14		Asr	1 A:	sp '	Ser	- G	lu G	11 <i>y</i> 50	Thr	· Me	et G	lu	Ar	9 Gl 15		u II	e Ly	s Asn	Cys 160
Se	יר ו	Phe	: A:	sn .	lie	16		hr	Ser	11	e A	rg	170		u Va	l Gl	n Lys	5 Glu 175	
Αţ	a I	Leu	Pł		80		's L	eu .	Asp	Va		a l 35	Pro	110	e Ası	p Asi	n Asr 190	n Asn	Thr
Se	r 1	ΊγΓ	Ar 19		.eu	11	e S	er (	Cys	As 20		16	Ser	Val	110	20:		Ala	Cys
Pr		ys 10	11	e S	er	Ph	e GI		Pro 215	110	e Pr	0	ile	His	7 7 7 7 2 2 C	•	. Ala	Pro	Ala
G1: 22:		he	Αl	a I	le	Le	u Ly 23		ys	Ası	n As	P	Lys	Thr 235		Asr	n Gly	Lys	Gly 240
Pro	<b>,</b> C	ys	Ly	s A	sn	Va   245		r T	hг	Val	Gl		Су <b>s</b> 250	Thr	His	Gly	ile	Arg 255	Pro
Val	. V	al	Se		hr 60	Glr	ı Le	u L	eu 		As . 25		Gly	Ser	Leu	Ala	Glu 270	Glu	Glu
Val	V	əl	1 l c 27:		9	Ser	As	p A		Phe 280		r	Asn	Asn	Ala	Lys 285		Ile	11e
Val		l n 20	Lev	4 Ly	/5	Glu	Se		a l 95	Glu	П	e /	Asn	Cys	Thr 300		Pro	Asn	Asn
Asn 305		1	Arg	, Ly	/S :	Ser	310	• H:	i s	lle	GL	y 1	Pro	Gly 315	Arg	Ala	Phe	Tyr	Thr 320
Chr	Gl	<b>y</b> :	Glu	11		1 l e 325	Gly	/ As	s <b>p</b>	Ile	Arg		31n 330	Ala	His	Cys	Asn	I le 335	Ser
				34	0						345	•					350	Leu	
			555						-	360						365		Gly	
	37	0						37	5						380			Phe	
9r 85	СУ	s /	Isn	Se	rT	hr.	390		u F	he	Asn	<b>.</b> S		Thr 3 <b>9</b> 5	Trp	Asn	Asn	Asn	Thr 400
					4	05				•		4	10					Cys 415	_
				420	)						425						430	Met	
		4	35						4	40						445		Thr	
	450	)						45	5						460			Glu	
<b>5</b> 5				,,			470						•	475				Glu	480
Yr	Lys	<b>.</b> T	уr	Lys		8 L 8 S	Val	Ly	s 1	le	Glu		го ( 90	Leu	Gly	Val	Ala	Pro 4 <b>9</b> 5	Thr
/S	Ala	L		Arg 500		<b>'9</b>	Vel	Val	G		Arg 505	G	tu i	LYS					

(	2) [	NFOR	MATI	ON	FOR S	1 P3	D NO	:17:								
		(i)	(A) (B) (C)	LEM TYP STR	IGTH: PE: n	148 IUC L E DNES	4 bas ic as S: s	ingle	airs							•
		! <b>( )</b>	MOLE	CULE	TYP	E: D	NA (g	genor	nic)							
	(	ix)	(B)	NAM LOC	E/KE' ATJOI ER II	V: 1.	. 147									
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:  ATG GAT GCA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG TGT GGA 4																
Me	G GA t As	T GC	A AT	G A/	AG AG /S Ar 5	iA GG 'g Gl	G CT y Le	C TGI	C TG S Cyt	s Va	G CTI	G CTG	CTG Leu	TGT Cys 15	Gly	48
GC.	A GT a va	C TT i Ph	e Va	T TC	יר אר 20 סי	C AG	C EAU	G GA/ n Glu 25	j He	CA1	F GCC	CGA Arg	TTC Phe 30	AFS	AGA Arg	96
GG:	C GC y Al	C AG a Ar 3	g Th	A GA r Gl	A AA U Ly	A TTI	G TG( U Trp 40	G GTC D Val	The	GTC Val	TAT	TAT Tyr 45	Gly	GTA Val	CCT Pro	144
GT( Val	S TGI L Tri 50	Ly	G GA	A GC J Al	A ACI a Thi	C ACC	Thr	Leu	TTT Phe	Cys	GCA Ala 60	Ser	GAT Asp	GCT Ala	AAA Lys	192
GCA Ala 65	Tyr	GA'	T ACA	GAI	G GT/ u Val 70	His	AAT Asn	GTT Val	TGG	GCC Ala 75	Thr	CAT	GCC Ala	TGT Cys	GTA Val 80	240
Pro	ACA Thr	GAC ASP	Pro	AAC ASF 85	1 Pro	Gln	GAA Glu	GTA Val	GTA Val 90	TTG Leu	GTA Val	AAT Asn	GTG Val	ACA Thr 95	GAA Glu	288
Asn	Phe	Asr	100	1 mg	Lys	Asn	Asp	ATG Met 105	Val	Glu	Gln	Met	His 110	Glu	Asp	336
ile	Ile	Ser 115	Leu	1rp	Asp	Gln	Ser 120	CTA Leu	Lys	Pro	Cys	Val 125	Lys	Leu	Thr	384
CCA Pro	CTC Leu 130	TGT Cys	GTT Val	AGT Ser	TTA Leu	AAG Lys 135	TGC Cys	ACT Thr	GAT ASP	TTG Leu	GGG Gly 140	AAT Asn	GCT Ala	ACT Thr	AAT Asn	432
ACC Thr 145	AAT Asn	AGT Ser	AGT Ser	AAT Asn	ACC Thr 150	AAT Asn	AGT Ser	AGT Ser	AGC Ser	GGG Gly 155	GAA Glu	ATG Met	ATG Met	ATG Met	GAG Glu 160	480
AAA Lys	GGA Gly	GAG Glu	ATA Ile	AAA Lys 165	AAC Asn	TGC Cys	TCT Ser	TTC Phe	AAT Asn 170	ATC Ile	AGC Ser	ACA Thr	AGC Ser	ATA Ile 175	AGA Arg	528
GGT Gly	AAG Lys	GTG Val	CAG Gln 180	AAA Lys	GAA Glu	TAT Tyr	GCA Ala	777 Phe 185	TTT Phe	TAT Tyr	AAA Lys	CTT L <b>eu</b>	GAT Asp 190	ATA 1 i e	ATA I i e	576

CCA ATA GAT AAT GAT ACT ACC AGC TAT ACG TTG ACA AGT TGT AAC ACC Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

624

	TCA Ser	GT Va 21	l i	ATT Ile	ACA Thr	Glr	G GC	a C	gt ( ys (	Pro	AA( Lys	5 G1 5 Va	A T	er P	111 Phe 220	GAG	CC/	A ATT	CCC Pro	672
	11e 225	CA'	5 1	AT yr	TGT Cys	GCC Ala	23	0 A	CT C	igt ily	TTT Phe	GC Al	G A' a ! 23	le L	TA eu	AAA Lys	TGT	AAT Asr	AAT Asn 240	720
A	AG ys	ACC Thr	P	TC /	AAT Asn	GGA Gly 245	AC. Th	A GI	GA C	CA '	TGT Cys	AC Th 25	r As	IT G	TC	AGC Ser	ACA Thr	GTA Val 255	GLN	768
T C	GT YS	ACA Thr	H	is (	GGA Gly 260	ATT	AGO	CC Pr	A G	al V	51A /e l 265	TC/ Se	A AC	T C/ r GI	AA ln	CTG Leu	CTG Leu 270	TTG Leu	AAT Asn	816
G	ly :	Ser	27	14 A	lat	Glu	Glu	Gl	u Va 28	al v 30	al	Ιle	Ar	g Se	9.	A i a 285	Asn	TTC Phe	Thr	864
As	ip /	15n 290	AL	al	ys T	hr	Ile	1 l	e Va 5	il G	ln	Leu	AS	n Gl 30	n :	Ser	Val	GAA Glu	Ile	912
30	5	ys	Τh	r G	ly A	la i	51y 310	His	s Cy	s A:	sn	Ιίe	Sei 319	Ar	9 /	Ala	Lys	TGG Trp	Asn 320	960
AL	a T	hr	Lei	u Ly	/s G 3:	ln 1 25	le	Ala	) Se	r Ly	ys	1 eu 330	Arg	GL	u (	iln	Phe	GGA Gly 335	Asn	1008
ASI	n L	ys `	Ihr	34	e II	le P	he	LYS	Gli	n Se 34	5	Ser	GLY	GE	y #	\sp	Pro 350	GAA Glu	Ile	1056
Va		3	55	Se	r Ph	ie A	sn	Cys	360	/ GI	у (	Slu	Phe	Pho	e 1	уг 65	Cys	AAT Asn	Ser	1104
Int	37	0 L	eu	Pho	P AS	n S	er :	1hr 375	Trp	Ph	e A	lsn	Ser	Th:	ר ז ס	ГÞ	Ser	ACT Thr	Glu	1152
385	` 5e	r A	sn	ASF	חו ו	39	20 20	Gly	Ser	As	p ī	hr	1 i e 395	Thr	· L	eu i	Pro		Arg 400	1200
Ile	Ly	s G	ln	Phe	405	2 As	in P	let	Trp	Gli	n G 4	10	Vel	Gly	Ļ	ys i	Ala	ATG Met 415	Tyr	1248
Ala	Pr	o Pi	0	1 le 420	Ser	r Gi	y G	iln	Ite	Arg 425	9 C	ys	Ser	Ser	• 🗚	sn i	11e 430	ACA Thr	GLY	1296
.eu	Let	43	5	Thr	Arg	) A8	ÞG	ly	Gly 440	Asr	7 A:	sn .	Asn	Asn	4 4	ly : 45	Ser	GAG Glu	Ile	1344
he	Arg 450	Pr	0	Gly "	GIA	Gi	y A 4	SP   55	Met	Arg	) A:	S.P.	Asn	1rp 460	A	rg :	Ser	GAA Glu	Leu	1392
yr 65	Lys	Ty	r 1	Lys	Val	Va 47	l L.	ys	Ile	Glu	ı Pı	ro (	.75	Gly	V	∌i /	Nia	CCC Pro	ACC Thr 480	1440
AG	GCA	**	G /	AGA	AGA	GT	G	rg (	CAG	AGA	G/	W /	w	T G	AGI	CGGC	ccc			1484

Lys Ala Lys Arg Arg Val Val Gin Arg Glu Lys 485 490

### (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 491 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 10 15

Ala Val Phe Val Ser Pro Ser Gin Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

Ata Tyr Asp Thr Glu Val His Asn Val Trp Ata Thr His Ata Cys Val
65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gin Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu
145 150 155 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro 210 215 220

Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Île Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn 260 265 270

Gly Ser Leu Ala Glu Glu Glu Val Val 11e Arg Ser Ala Asn Phe Thr 275 280 285

As	290		a Lys	; IN	r 11	295			ı Let	J AST	300		val	GLU	lle	
As:		s Thr	- Gly	Al	a Gl 31	y His O	Cy:	s Asr	ı Ile	Ser 315		Ala	Lys	Trp	Asn 320	
Αla	3 Thi	Lec	ı Lys	Gt/ 32:		e Ala	Ser	Lys	330		GLU	Gln	Phe	Gly 335	Asn	
Asr	ı Lys	Thr	11e 340		Ph:	e Lys	Gtr	345		Gly	Gly	Asp	Pro 350		Ile	
Val	Thr	His 355		Phe	. Asr	ı Cys	Gly 360		Glu	Phe	Phe	1 yr 365	Cys	Asn	Ser	
Thr	Glr 370	-	Phe	Asn	Ser	Thr 375	Trp	Phe	Asn	Ser	Thr 380		Ser	Thr	Glu	
Gly 385		Asn	Asn	Thr	G11	Gly	Ser	Asp	Thr	1 l e 395		Leu	Pro	Cys	Arg 400	
Ile	Lys	Gln	Phe	1 l e 405		Het	Trp	Gln	Glu 410		Gly	Lys	Ala	<b>Het</b> 415	Tyr	
Ala	Pro	Pro	1 l e 420	Ser	Gly	Gln	lle	Arg 425	Cys	Ser	Ser	Asn	11e 430	Thr	Gly	
Leu	Leu	Leu 435	Thr	Arg	Asp	Gly	Gly 440		Asn	Asn	Asn	Gly 445	Ser	Glu	Ile	
Phe	Arg 450	Pro	Gly	Gly	Gly	Asp 455	Het	Arg	Asp	Asn	1rp 460	Arg	Ser	Glu	Leu	
465		-			470	Lys				475	Gly	Val	Ala	Pro	Thr 480	
Lys	Ala	Lys		Arg 485	Væl	Val	Gln	Arg	490	Lys						
(2) INFORMATION FOR SEG ID NO:19:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1448 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: DNA (genomic)  (ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 11439 (D) OTHER INFORMATION:																
ATC						PT I O GGG						CTG	CTG	TGT	GGA	48
						Gly										***
						AGC   Ser								Arg		96
GGC Gly	GGC	AGA ( Arg ( 35	GTA (	iAA i	AAG Lys	TTG :	TGG Trp 40	GTC Val	ACA Thr	GTC Val	TAT Tyr	TAT Tyr 45	GGG Gly	GTA Val	CCT Pro	144

		P L)				CC AC						Ser				. 192
	a Ty				lu V	TA CA ol Hi 70					Thr					240
				O A:		A CA				Leu						288
				t Tr		A AA' S ASI			Val							336
			r Le			T CAI		Leu								384
		Cys				A AA1 u Asr 135	Cys									432
	Asn					A ACG										480
					r Thi	A AGC										528
Ala	Leu	Phe	180	- Ly! )	i Lei	T GAT J Asp	Vəl	Val 185	Pro	ile	Asp	Asn	Asn 190	Asn	Thr	576
Ser	Туг	Arg 195	Leu	iIle	: Ser	tgt Cys	Asp 200	Thr	Ser	Val	Ile	Thr 205	Gln	Ala	Cys	624
Pro	Lys 210	lle	Ser	Phe	Git	Pro 215	lle	Pro	ile	His	Tyr 220	Cys	Ala	Pro	AL	672
Gly 225	Phe	Ala	Ile	Leu	230		Asn	Asp	Lys	1hr 235	Phe	Asn	Gly	Lys	Gly 240	720
Pro	Cys	Lys	Asn	Val 245	Ser	ACA Thr	Val	Gln	Cys 250	Thr	His	Gly	Ile	Arg 255	Pro	768
Val	Val	Ser	Thr 260	Gln	Leu	Leu	Leu	Asn 265	Gly	Ser	Leu	Ala	Glu 270	Glu	Glu	816
Val	Val	1 l e 275	Arg	Ser	Asp	AAT	Phe 280	Thr	Asn	Asn	Ala	Lys 285	Thr	Ile	Ile	912
Val	Gln 290	Leu	Lys 	Glu	Ser	GTA Val 295	Glu	lle	Asn	Cys	1hr 300	Gly	Ala	Gly	His	
Cys 305	Asn	Ite	Ser	Arg	Ala 310	Lys	Trp	Asn	Asp	1hr 315	Leu	Lys	Gln	Ile	720	960
AIA	***	LIA	AuA	LLAK.		TTT	<b>4</b>	~~ i	~~~	ALA	~!^	g i L		~41	-46	1008

11	e Ly	ys L	eu #	\rg	325	Gli	n Ph	e Gl	u As	n Ly 33		r Ild	e Val	Phe	335		
TC Se	C TC	A G	ly G	igg ily i+0	GAC ASP	CC# Pro	GA/	A AT	T GT/ e Va 349	l Me	G CAC	C AG1	TTT Phe	AAT ASD 350	Cys	GGA Gly	1056
GG/ GL)	A GA Y Gi	A T1 u Ph 35	e P	TC he	TAC Tyr	TGT Cys	AAT ASF	7 C# 360	A ACA	CAA Glr	CTC	i IT1 I Phe	AAT Asn 365	Ser	ACT Thr	TGG Trp	1104
Asr	AA A ASI - 370	n As	T A	CT hr	GAA Glu	GGG Gly	TCA Ser 375	Asn	AAC Asn	ACT Thr	GAA Glu	GGA Gly 380	Asn	ACT	ATC	ACA Thr	1152
CTC Leu 385	Pro	A TG	C AC	3A /	Ιle	AAA Lys 390	CAA Gln	ATT 1le	ATA	AAC Asn	ATG Met 395	TGG Trp	CAG	<b>GAA</b> Glu	GTA Val	GGA Gly 400	1200
AAA Lys	GCA	ATI Me	G TA	r	CC Na 1	CCT Pro	CCC Pro	ATC Ile	AGA Arg	GGA Gly 410	CAA Gln	ATT	AGA Arg	TGT Cys	TCA Ser 415	TCA Ser	1248
AAT Asn	ATT	ACA Thr	GG G1 42	y L	TG (	CTA Leu	TTA Leu	ACA Thr	AGA Arg 425	GAT Asp	GGT Gly	GGT Gly	ÄTT Ile	AAT Asn 430	GAG Glu	AAT Asn	1296
GGG Gly	ACC Thr	GAG Glu 435	11	C T	TC A	IGA Irg	Pro	GGA Gly 440.	GGA Gly	GGA Gly	GAT Asp	ATG Met	AGG Arg 445	GAC Asp	AAT Asn	TGG Trp	1344
Arg	AGT Ser 450	GAA Glu	TT/	A T	AT A yr L	ys '	TAT . Tyr 455	AAA Lys	GTA Val	GTA Val	Lys	ATT Ile 460	GAA Glu	CCA Pro	TTA Leu	GGA Gly	1392
STA Val 465	GCA Ala	CCC Pro	ACC	L)	/S A	CA / la i 70	AG A	AGA Arg	AGA Arg	Val	GTG Val 475	CAA Gln	AGA Arg	GAA Glu	AAA Lys	TG	1439
\GCG	GCCG	C															1448

### (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 479 amino acids (B) TYPE: amino acid

  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys
50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu

					85	5				9	0				95	
His	s Ph	e As		et 00	Trp	) Ly	s As	n As	n Me 10:		l Git	. Glr	Het	Gln 110		Asp
ш	e Il	e Se 11		eu	Trp	) Ası	p Gl	n Se 12		u Ly:	s Pro	Cys	Val 125		Leu	Thr
Pro	13:		s V	al	Thr	Lev	J ASI 13!		s Ly:	s Ast	o Val	Asn 140		Thr	Asn	Thr
Thr 145		n As	p Se	er	Glu	150		r Me1	t <sub>,</sub> Gli	J Arg	155		He	Lys	Asn	Cys 160
					165					170				-	175	•
Ala	Leu	ı Ph		/r (	Lys	Leu	ASP	vat	185		ile	Asp	Asn	190	Asn	Thr
	·	19	5				-	200	)		Val		205			-
	210	)					215				His	220				
225						230			·	·	Thr 235			_	·	240
	·	·		2	45					250			•		255	
			26	D					265	·	Ser			270		
		275						280			Cys		285			
	290						295				Thr	300	•		•	
305						310	·	·		·	315 Thr		•			320
	•			3	25					330	His				335	
			340	)					345		Leu			350		
•		355				•		360			Glu		365			•
	370						375					380				
385					3	<b>390</b>					395 Gln					400
				40	)5					410	Gly				415	
			420						425		Asp			430		
- ` ,		435			- •			440		,			445			

455 450 val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gin Arg Glu Lys 470 (2) INFORMATION FOR SEQ ID NO:21: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1484 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: DNA (genomic) (ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 1..1454 (D) OTHER INFORMATION: (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21: ATG GAT GCA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG CTG TGT GGA 48 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly

1 5 10 GCA GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30 GGC GCC AGA ACA GAA AAA ITG TGG GTC ACA GTC TAT TAT GGG GTA CCT Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro GTG TGG AAG GAA GCA ACC ACC ACT CTA TTT TGT GCA TCA GAT GCT AAA 192 Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys GCA TAT GAT ACA GAG GTA CAT AAT GTT TGG GCC ACA CAT GCC TGT GTA 240 Ala Tyr Asp Thr Glu Val His Ash Val Trp Ala Thr His Ala Cys Val CCC ACA GAC CCC AAC CCA CAA GAA GTA GTA TTG GTA AAT GTG ACA GAA Pro Thr Asp Pro Asn Pro Gin Giu Vai Vai Leu Vai Asn Vai Thr Giu 288 AAT TIT AAC ATG TGG AAA AAT GAE ATG GTA GAA CAG ATG CAT GAG GAT Asn Phe Asn Het Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp ATA ATC AGT ITA TGG GAT CAA AGC CTA AAG CCA TGT GTA AAA TTA ACC 384 The The Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr CCA CTC TGT GTT AGT TTA AAG TGC ACT GAT TTG GGG AAT GCT ACT AAT 432 Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn ACC AAT AGT AGT AAT ACC AAT AGT AGT AGC GGG GAA ATG ATG ATG GAG 480 Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu AAA GGA GAG ATA AAA AAC TGC TCT TTC AAT ATC AGC ACA AGC ATA AGA 52B Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg

GGT AAG GTG CAG AAA GAA TAT GCA TTY TTT TAT AAA CTT GAT ATA ATA

Gly Lys Val Gin Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp lie lie

					180	)				18	15				190	)		
C:	CA	AT.	e As	AT /	AAT Asn	GA' Asi	T AC	T AC	C AC	r Ty	T AC	G TT(	ACA J Thr	Ser 205	Cys	AAC Asn	ACC Thr	624
	:r		H						's Pr		G GT/ s Val			Glu				672
	e							o Al			T GCC e Ala		Leu					720
							Thi				7 ACA 5 Thr 250	Asn						768
TG Cy	T /	ACA Thr	CA Hi	s G	GA ly 60	ATT	AGG	CC/	A GT	4 GTA 1 Val 265	TCA Ser	ACT	CAA Gln	CTG Leu	CTG Leu 270	TTG L <b>e</b> u	AAT Asn	816
				J A						Val	ile							954
GA(	<b>A</b> C	AT Asn 290	GC1 Ala	T AA	VA /S	ACC Thr	ATA	11e	. Val	Gln	CTG Leu	AAC Asn	CAA Gin 300	TCT Ser	GTA Val	GAA Glu	ATT Ile	912
AA1 Asr 305	) C	GT YS	ACA	G G	) Y	GCT Ala	GGA Gly 310	His	Cys	AAC Asn	ATT	AGT Ser 315	AGA Arg	GCA Ala	AAA Lys	TGG Trp	AAT Asn 320	960
GCC Ala	. A	CT hr	TTA Leu	. AA	<b>s</b> (	AG iln i25	ATA Ile	GCT Ala	AGC Ser	Lys	TTA Leu 330	AGA Arg	GAA Glu	CAA Gln	TTT Phe	GGA Gly 335	AAT Asn	1008
					e i						TCA Ser							1056
GTA Val	AI TI	hr	CAC His 355	AG Se	T T	TT .	AAT Asn	TGT Cys	GGA Gly 360	GGG Gly	GAA Glu	TTT Phe	TTC Phe	TAC Tyr 365	TGT Cys	AAT Asn	TCA Ser	1104
ACA Thr	G	Ln 70	CTG Leu	TT'	T A	AT /	AGT Ser	ACT Thr 375	TGG Trp	TTT Phe	AAT Asn	AGT Ser	ACT Thr 380	TGG Trp	AGT Ser	ACT Thr	GAA Glu	1152
GGG GLy 585	TO Se	EA A	AAT Asn	ASI	C A	hr (	GAA Glu 390	GGA Gly	AGT Ser	GAC Asp	ACA Thr	ATC Ile 395	ACA Thr	CTC Leu	CCA Pro	TGC Cys	AGA Arg 400	1200
le	AA Ly	VA (	CAA Gln	TT1 Phe	: 1	TA / le /	MC Nan	ATG Met	GTG Val	CAG Gln	GAA Glu 410	GTA Val	GGA	AAA Lys	GCA Ala	ATG Met 415	TAT Tyr	1248
icc	CC Pr	T C	Pro	ATC   le 420	S	GC 6 er 6	igA ily	CAA Gln	ATT	AGA Arg 425	TGT . Cys	TCA Ser	TCA Ser	AAT Asn	ATT 11e 430	ACA Thr	GGG Gly	12 <del>96</del>
		u L									AAC Asn		Asn					1344
		g P					ly i				GAC Asp	Asn						1392

TAT AAA TAT AAA GTA GTA AAA ATT GAA CCA TTA GGA GTA GCA CCC ACC Tyr Lys Tyr Lys Val Val Lys lie Glu Pro Leu Giy Val Ala Pro Thr 465 470 475 480

AAG GCA AAG AGA AG AGTGGTGCAG AGAGAAAAAT GAGCGGCCGC Lys Ala Lys Arg

1484

#### (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 484 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Cys Gly
1 5 10 15

Ata Val Phe Val Ser Pro Ser Gln Glu Ite His Ata Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro
35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys
50 60

Ala Tyr Asp Thr Glu Val His Ash Val Trp Ala Thr His Ala Cys Val 65 70 75 80

. Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu 145 150 155 160

Lys Gly Glu Ite Lys Asn Cys Ser Phe Asn Ite Ser Thr Ser Ite Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ile Thr Gin Ala Cys Pro Lys Val Ser Phe Giu Pro Ile Pro 210 215 220

lle His Tyr Cys Ala Pro Ala Cty Phe Ala lle Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn

													7	5	
			26	0				26	5				270	1	
٥١	y Se	r Le 27	eu Al	a Gi	u Gl	u Gli	اه۷ د 280		l Ile	e Arg	Ser	Alá 285		Phe	Thr
Ası	P As 29		a Ly	s Th	r 11	e 116 295		Glr	n Lei	J ASI	300		Val	Glu	ılle
Asi 305		s Th	r Gl	y Al	a Gl		Cys	ASF	ı ile	315		Ala	Lys	īrp	Asn 320
Ala	Thi	Le	u Ly:	s Gl/ 325	n Ile	e Ala	Ser	Lys	330		Glu	Gln	Phe	Gly 335	
Asr	Lys	Th	r Ile 340	e Ile	Phe	. Lys	Gln	Ser 345		Gly	Gly	Asp	Pro 350	Glu	Ile
Val	Thr	#i: 35!	s Ser	Phe	Asr	Cys	Gly 360	Gly	Glu	Phe	Phe	1yr 365	Cys	Asn	Ser
Thr	Gln 370	Lei	, Phe	: Asn	Ser	Thr 375	Trp	Phe	Asn	Ser	Thr 380	Trp	Ser	Thr	Glu
Gly 3 <b>85</b>	Ser	Asr	ASP	Thr	Glu 390	Gly	Ser	Asp	Thr	11e 395	Thr	Leu	Pro	îys	Arg 400
He	Lys	Glr	Phe	11e 405	Asn	Met	Val	Gln	Glu 410	Val	Gly	Lys	Ala	<b>Ket</b> 415	Tyr
Ala	Pro	Pro	11e 420	Ser	Gly	Gln	Ile	Arg 425	Cys	Ser	Ser	Asn	11e 430	Thr	Gly
Leu	Leu	L eu 435	Thr	Arg	Asp	Gly	Gly 440	Asn	Asn	Asn		Gly 445	Ser	Glu	Ile
Phe	Arg 450	Pro	Gly	Gly	Gly	Asp 455	Net .	Arg	Asp		Trp 460	Arg	Ser	Glu	Leu
Tyr 465	Lys	Tyr	Lys	Val	Val 470	Lys	Ile	Glu	Pro	Leu 475	Gly	Val	Ala	Pro	Thr 480

## (2) INFORMATION FOR SEG ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1448 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:

Lys Ala Lys Arg

- (A) NAME/KEY: CDS (B) LOCATION: 1..1438
- (D) OTHER INFORMATION:

#### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ATG GAT GEA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG TGT GGA Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10

GCA GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA ALA VAL Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25

48

GG	C GG( y Gl;	C AGA	g Va	A GAA	Lys	i TTG i Leu	TGG Trp 40	Val	ACA Thr	Val	TAT	TAT Tyr 45	GGG Gly	GTA Val	CCT Pro	144
GT Va	G TG( L Tr; 5(	Lys	A GAA	A GCA J Ala	ACC Thr	ACC Thr	Thr	CTA	TTT Phe	TGT Cys	GCA Ala 60	TCA Ser	GAT Asp	GCT Ala	AAA Lys	192
GC/ A L &	a Tyr	GAT ASE	ACA Thr	GAG	GTA Val 70	CAT His	AAT Asn	GTT Val	TGG	GCC Ala 75	ACA Thr	CAT His	GCC	TGT Cys	GTA Val 80	240
CCC Pro	ACA Thr	GAC Asp	CCC Pro	AAC Asn 85	Pro	CAA	GAA Glu	GTA Val	GTA Val 90	Leu	GAA Glu	AAT Asn	GTA Val	ACA Thr 95	GAA Glu	288
CA1 His	TTT Phe	AAC Asn	ATG Met 100	Trp	AAA Lys	AAT Asn	AAC Asn	ATG Met 105	GTA Val	GAA Glu	CAG Gln	ATG Met	CAG Gln 110	GAG Glu	GAT ASP	336
						CAA Gln										384
		Cys				AAT Asn 135										432
	Asn					ACG Thr										480
						AGC Ser										528
						GAT ASP										576
er	Tyr	Arg 195	Leu	Ite	Ser		Asp 200	Thr	Ser	Val	Ile	Thr 205	Gin	Ala	Cys	624
ro	Lys 210	Ile	Ser	Phe	Glu	Pro 215	Ile	Pro	Ile	His	Tyr 220	Cys	Ala	Pro	Ala	672
l y 25	Phe	Ala	lle	Leu	Lys 230	TGT Cys	Asn	Asp	Lys	Thr 235	Phe	Asn	Gly	Lys	Gly 240	720
ro	Cys	Lys	Asn	Val : 245	Ser	ACA Thr	Val	Gin	Сув 250	Thr	His	Gly	ile	Arg 255	Pro	764
al	Val	Ser	Thr 260	Gin I	Leu	CTG :	Leu	A <b>s</b> n 265	Gly	Ser	Leu	Ala	61u 270	Glu	Glu	810
al	Val	1 le 1 275	Arg :	Ser /	Asp		Phe 280	Thr	Asn	Asn	Ala	L ys 285	Thr	Ile	ile	86
					Ser	GTA Val 295										91
CT	244	ATT	AGT	AGA 4	GCA		TGG	AAT	GAC	ACT	TTA	AAA	CAG	ATA	GTT	96

Cys 305		ile	Ser	Arg	Ala 310		Тгр	Asn	Asp	Thr 315	Leu	Lys	Gln	He	Val 320	
ATA !le	AAA Lys	TTA Leu	AGA Arg	GAA Glu 325	CAA	TTT Phe	GAG Glu	AAT Asn	AAA Lys 330	ACA Thr	ATA Ile	GTC Val	TTT Phe	AAT Asn 335	CAC His	1008.
TCC Ser	TCA Ser	GGA Gly	GGG Gly 340	GAC Asp	CCA Pro	GAA Glu	ATT	GTA Val 345	ATG Met	CAC	AGT Ser	TTT Phe	AAT Asn 350	TGT Cys	GGA Gly	1056
GGA Gly	GAA Glu	777 Phe 355	TTC Phe	TAC	TGT Cys	AAT Asn	TCA Ser 360	ACA Thr	CAA Gln	CTG Leu	TTT Phe	AAT Asn 365	AGT Ser	ACT Thr	TGG Trp	1104
AAT Asn	AAT Asn 370	AAT Asn	ACT Thr	GAA Glu	GGG Gly	TCA Ser 375	AAT Asn	AAC Asn	ACT Thr	GAA Glu	GGA Gly 380	AAT Asn	ACT Thr	ATC lie	ACA Thr	1152
CTC Leu 385	CCA Pro	TGC Cys	AGA Arg	ATA 1le	AAA Lys 390	CAA Gln	ATT	ATA Ile	AAC Asn	ATG Met 395	GTG Val	CAG Gln	GAA Glu	GTA Val	GGA Gly 400	1200
			TAT Tyr													1248
			GGG Gly 420													1296
			ATC Ile			Pro										1344
Arg			TTA Leu		Lys											1392
			ACC Thr	Lys											Т	1438
GAGC	GGCC	GC														1448

### (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 479 amino acids
  - (B) TYPE: amino scid (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Asp Ala Net Lys Arg Gly Leu Cys Cys' Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys
50 55 60

	a 1	yr	As	p T	hг	Glu	7 Va		is As	in Va	al Tr	P AL	_	His	Ala	Cys	Val 80
Pr	o t	hг	As	p P	ro	Asn 85	Pr	o Gl	n Gl	u Va	l Va		u Glu	AST	Val	Thr 95	Glu
Hi	s P	he	As		e t 00	Trp	Lys	s As	n As	n Me	t Va 5	l Gli	Gln	Met	Gln 110	Glu	Asp
It	e. I	le	Se 11		eu '	ľτp	Asp	Gli	n Se 12		u Ly:	s Pro	Cys	Val 125	Lys	Leu	Thr
Pr		eu 30	Cy:	s Vá	at 1	hr	l eu	139		s Ly	s Ast	val	Asn 140	Ala	Thr	Asn	Thr
Th:		sn	Ası	<b>5</b> 5€	פר 0	ilu	Gly 150		r Me	t Gl	u Arg	Gly 155		Ιle	Lys	Asn	Cys 160
Sei	- Ph	e	Asr	11		hr 65	Thr	Ser	· Ile	e Ar	9 Asp 170		Val	Gln	lys	Glu 175	Туг
Ala	Le	u	Phe	18		ys	Leu	Asp	Val	185	Pro	ile	Asp	Asn	Asn 190	Asn	Thr
Ser	Ту		Arg 1 <b>9</b> 5		u i	l e	Ser	Cys	200		Ser	Val	Iie	1hr 205	Gln	Ala	Cys
Pro	21		ite	Se	r Pl	he i	Glu	Pro 215		Pro	Ile	His	Tyr 220	Cys	Ata	Pro	Ala
31 y 225	Ph	e /	Nia	110	e Lo		230	Cys	Asn	Asp	Lys	Thr 235	Phe	Asn	Gly	Lys	Gly 240
					24	5					250			-		255	
al	Va			260	)					265		•			270		
	Val	2	75						280		Asn	•		285			
	290	)						295			Asn		300				
05						3	10				Asp	315					320
				,	32	5					Lys 330					335	
				340						345	Met			٠	350		•
		3:	55						360		Gln			365			
	370						3	375			Thr		380				
85						39	90				Asn	395					400
					405	i					Gly 410			-	·	415	
SN	ile	T		Gly 420			eu l	.eu		Arg 425	ASP	GLY	GLY		Asn 430		ASN

145

Gly	/ Thr	Gl:		e Ph	e Ar	g Pro	Gly 440		/ Gly	Asp	Met	Arg 445		Asn	Тгр	
Arg	Ser 450		ı Lei	ц Туг	r Lys	455		Val	Val	Ŀys	1 i e 460		Pro	l eu	Gly	
Vai 465		Pro	; Thi	r Ly:	5 Ala 470	e lys )	Arg	Arg	yat	Val 475	Gln	Arg	Glu	Lys		
(2)	INF	ORMA	T I O	N FOF	SEC	ID.	NO:2	5:								
	(i	(	A) L B) 1 C) 9	ENGT TYPE:	H: 1 DEDN	CTER 571 :leic !ESS: lin	base aci sin	pai d	rs							
	(ii	) MC	LECL	JLE T	YPE:	DNA	(ge	nomi	c)							
	(ix	) FE			KEY:	CDS										
		(	B) L	OCAT	ION:	1 ANGO		:								
		•	-, -													
	(xi	) SE	QUEN	CE D	ESCR	IPTIC	ON:	SEQ	ID NO	25:	:					
ATG	GAT	GCA	ATG	AAG	AGA	GGG Gly	CTC	TGC	TGT	GTG	CTG	CTG	CTG Leu	TGT	GGA GLV	48
1	vsh	~	net	5		J.,		-,-	10				•	15	,	
						AGC Ser										96
~(4	VO	riie	20		,,,		••••	25					30			
GGC	GCC	AGA	ACA	GAA	AAA	TTG Leu	TGG	GTC	AÇA	GTC	TAT	TAT	GGG	GTA	CCT	144
GLY	ALB	35		4.5	-,,		40	•••	••••	•••	.,.	45	4.,		•••	
						ACC Thr										192
Vat	50	LYS	3.0	~		55	,		7	Cys	60	JC.	лар	~.•	-,3	
GCA	TAT	GAT	ACA	GAG	GTA	CAT His	AAT	GTT	TGG	GCC	ACA	CAT	GCC	TGT	GTA	240
65	ıyr	ASP	Int	910	70	піъ	POLI	V-0.	11 p	75	••••	піз	~10	Cys	80	
CCC	ACA	GAC	CCC	AAC	CCA	CAA	GAA	GTA	GTA	TTG	GTA	AAT	GTG	ACA	GAA	288
Pro	inr	ASP	PFO	85 85	Pro	Gln	Glu	V.	90	Leu	Vat	ASII	Val	95	GLU	
AAT	111	AAC	ATG	TGG	***	AAT	GAC	ATG	GTA	GAA	CAG	ATG	CAT	GAG	GAT	336
Asn	Phe	Asn	100		Lys	Asn	ASP	105	Val	Glu	GLN	Met	110	ern	ASP	
ATA	ATC	AGT	TTA	TGG	GAT	CAA	AGC	CTA	AAG	CCA	TGT	GTA	***	TTA	ACC	384
Ite	lle	Ser 115	Leu	Trp	ASP	Gln	Ser 120	Leu	LYS.	Pro	Cys	Val 125	LYS	ren	TNC	
CCA	CTC	TGT	GTT	AGT	ATT	AAG	TGC	ACT	GAT	TTG	GGG	AAT	GCT	ACT	AAT	432
Pro	130	Cys	Val	Ser	Leu	Lys 135	Cys	Thr	Asp	Fen	140	Asn	ALB	Thr	Asn	

ACC AAT AGT AGT AAT ACC AAT AGT AGT AGC GGG GAA ATG ATG ATG TAGE THE ASE SET SET SET GLY GLU MET MET MET GLU 145 150 160

AAA GGA GAG ATA AAA AAC TGC TCT TTC AAT ATC AGC ACA AGC ATA AGA Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg

					16	5				17	0				175	i	
					Ly				la Pi		T TA				ile		576
CC. Pr	م م ا 0	le i	GAT 4sp 195	e e 1 Asr	GA As	T AC P Th	T A	00 40 nr Se 20	er Ty	AT AC	G TTI	G AC	A AGT r Ser 205	Cys	AAC	ACC	624
		et 1						s Pr			A TCO L Ser		Glu				672
	e Hi						o Al				G ATT B Ile 235	Leu					720
						Th					A AAT - Asn )						768
			is							Ser	ACT Thr						816
GGC Gly	AG'	L	75	GCA Ala	GAA Glu	GAA	GAC	GT/ Val 280	l Val	ATT	AGA Arg	TCT Ser	GCC Ala 285	AAT Asn	TTC Phe	ACA Thr	864
GAC ASP	AA1 Asr 290	A	a t	AA .ys	ACC Thr	ATA	ATA 11e 295	· Val	CAG Gln	CTG Leu	AAC Asn	CAA Gln 300	TCT Ser	GTA Val	GAA Glu	ATT lie	912
Asn 305	Cys	Th	r A	rg	Pro	310	Asn	Asr	1hr	Arg	AAA Lys 315	Ser	lie	Arg	ile	Gln 320	960
Arg	Gly	Pr	o G	ly /	17g 525	Ala	Phe	Val	Thr	11e 330	GGA Gly	Lys	He	Gly	Asn 335	Het	1008
Arg	Gln	AL	в Н 3	is ( 40	ys	Asn	Ile	Ser	Arg 345	Ala	AAA Lys	Trp	Asn	Ala 350	Thr	Leu	1056
.ys	Gin	355	e A 5	la S	er	Lys	Leu	Arg 360	Glu	Gln	TTT Phe	Gly	Asn 365	Asn	Lys	Thr	1104
lle :	1 l e 370	Pho	e Ly	ys G	in:	Ser	Ser 375	Gly	Gly	Asp	Pro	Glu 380	Ite	Val	Thr	His	1152
ier F 185	he	Asr	ı C)	rs G	ly (	11y 190	Glu	Phe	Phe	Туг	TGT Cys 395	Asn	Ser	Thr	Gln	Leu 400	1200
he A	lsn	Ser	Th	r T	rp F 05	he .	Asn	Ser	Thr	1rp 410	AGT Ser	Thr	Glu	Gly	Ser 415	Asn	1248
sn 1	hr	Glu	42 42	y S(	er A	\sp	Thr	Ile	Thr 425	Leu	Pro	Cys	Arg	1 le 430	Lys	Gln	1296
TT A	le	AAC Asn 435	He	G G	rg c	AG I	GAA Glu	GTA Val 440	GGA Gly	AAA Lys	GCA Ala	ATG Met	TAT Tyr 445	GCC Ala	CCT Pro	CCC Pro	1344

AGC Ser 450	Gly									1392
 AGA Arg	_							_	_	 1440
 GGA Gly										1488
 GTA Val							_			1536
 AGA Arg				TGA	GCG	G CC	GC			1571

#### (2) INFORMATION FOR SEG ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 522 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20  $\phantom{-}25\phantom{+}30\phantom{+}$ 

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro  $35 \hspace{1cm} 40 \hspace{1cm} 45$ 

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

Ata Tyr Asp Thr Glu Val His Asn Val Trp Ata Thr His Ata Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lya Cya Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu 145 150 155 160

Lys Gly Glu Ile Lys Acm Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Ash Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Ash Thr

		19	95				200	)				205			
Sei	r va 21		e Th	r Gl	n Ala	215		Lys	val	Ser	220		Pro	Ile	Pro
225		s Ту	r Cy	s Al	a Pro 230		Gly	Phe	Ala	235		Lys	Cys	Asn	Asn 240
Lys	5 Thi	- Ph	e As	n Gly 245		Gly	Pro	Cys	1hr 250		Val	Ser	Thr	Val 255	Gln
Cys	Thr	· Hi	s Gl		Arg	Pro	Val	Val 265	Ser	Thr	Gln	Ļeu	270	Leu	Asn
Gly	/ Ser	27:	u Ala 5	s Glu	GLU	Glu	val 280		110	Arg	Ser	Ala 285	Asn	Phe	Thr
Asp	290		a Lys	s Thr	Ile	I l e 2 <b>95</b>	Val	Gln	Leu	Asn	Gl n 300	Ser	Val	Glu	lle
Asn 305	•	Thi	r Arg	Pro	310	Asn	Asn	Thr	Arg	Lys 315	Ser	Ile	Arg	He	Gln 320
Arg	Gly	Pro	Gly	7 Arg 325		Phe	Val	Thr	11e 330	Gly	Lys	Ile	Gly	Asn 335	Het
Arg	Gln	Ala	340		Asn	I l e	Ser	Arg 345	Ala	Lys	Trp	Asn	Ala 350	Thr	Leu
Lys	Gln	11e 355	Ala	Ser	Lys	Leu	Arg 360	Glu	Gln	Phe	Gly	Asn 365	Asn	Lys	Thr
ile	1 le 370	Phe	Lys	Gln	Ser	Ser 375	Gly	Gly	Asp	Pro	Gl u 380	ile	Val	Thr	His
385			Cys	-	390				·	395					400
Phe	Asn	Ser	Thr	1 rp 405	Phe	Asn	Ser	Thr	1rp 410	Ser	Thr	Glu	Gly	Ser 415	Asn
			420					425					430	•	
		435	Met				440	-	·			445			
	450	·	Gln		_	455					460				
465	-	·	Gly		470				·	475				Ī	480
·	·	·	Asp	485		·			490				·	495	
Lys	Val	Val	Lys 500	He	Glu	Pro		Gly 505	Val	Ala	Pro	Thr	Lys 510	Ala	Lys

(2) INFORMATION FOR SEQ ID NO:27:

Arg Arg Val Val Gln Arg Glu Lys 515 520

- (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 1532 base pairs

  - (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

### (ix) FEATURE:

- (A) NAME/KEY: CDS
  (B) LOCATION: 1..1522
  (D) OTHER INFORMATION:

(xi)	SEQUENCE	DESCRIPTION:	SEQ	ID	NO:27:
------	----------	--------------	-----	----	--------

	( , ,	, 36	. 4021				<b>U</b>	J		U.L.	•					
	Asp				Arg				TGT Cys 10	Val						48
				Ser					ATC Ile							96
			Val					Val	ACA Thr							144
GTG Val	TGG Trp 50	Lys	GAA Glu	GCA Ala	ACC Thr	ACE Thr 55	ACT Thr	CTA Leu	TŢŢ Phe	TGT Cys	GCA Ala 60	TCA Ser	GAT ASP	GCT Alä	AAA Lys	192
	Tyr					His			TGG Trp							240
									GTA Val 90							288
									GTA Val							336
									AAG Lys							384
									GAT Asp							432
				Glu					AGA Arg							480
									GAT ASP 170							528
							Val		CCA Pro							576
						Cys			TCA Ser							624
Pro					GLU				ATA Ile							672
				Leu					AAG Lys						GGA Gly 240	720

			A AA1 s Asi		Ser					ihr						768
			A AC1 r Thr 260	Gir					GLY							816
			AGA Arg					Thr								864
		Leu	AAA Jelys													912
	Thr		AAA Lys													960
			ATA Ile													1008
			TGG Trp 340				Leu									1056
Glu	Gln	Phe 355	GAG Glu	Asn	Lys	Thr	1 l e 360	Val	Phe	Asn	His	Ser 365	Ser	Gly	Gly	1104
Asp	Pro 370	Glu	ATT	Val	Met :	His : 375	Ser	Phe	Asn	Cys	Gly 380	Gly	Glu	Phe	Phe	1152
1yr 3 <b>85</b>	Cys	Asn	TCA Ser	Thr	Gln ( 390	Leu I	Phe	Asn	Ser	Thr 395	Trp	Asn	Asn	Asn	Thr 400	1200
			AAT Asn					Asn					Pro			1248
ile	Lys	Gln	ATT I Ile 420	lie /	Asn A	tet \	al (	Gln 425	Glu	Val	Gly	Lys	Ala 430	Het	Tyr	1296
Ala	Pro	Pro 435	ATC /	Arg (	ily (	iln I 4	le /	Arg	Cys	\$er	Ser	Asn 445	He	Thr	Gly	1344
Leu	Leu 450	Leu	ACA /	Arg A	ASP G	ily G 55	ily I	ile i	Asn	Glu	Asn 460	Gly	Thr	Glu	Ile	1392
Phe . 465	Arg	Pro (	GGA (	Sty G 4	19 A 70	sp M	et /	lrg (	ASP	Asn 475	Trp	Arg	Ser	Glu	Leu 480	. 1440
TAT .	AAA Lys	TAT A	AAA G Lys V 4	TA G /al V .85	TA A	AA A ys I	TT (	ilu I	Pro 1	TTA   Leu	GGA Gly	GTA Val	GCA Ala	CCC Pro 495	ACC Thr	1488
		Lys /	AGA A Arg A 500				in A				T GA	GCGG	CCGC	:		1532

- (2) INFORMATION FOR SEQ ID NO:28:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 507 amino acids
    - (B) TYPE: amino acid (D) TOPOLOGY: linear
  - ·
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

At a Val Phe Val Ser Pro Ser Gln Glu Ile His Ata Arg Phe Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro
35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

Ata Tyr Asp Thr Glu Val His Asn val Trp Ata Thr His Ata Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gin Glu Val Val Leu Glu Asn Val Thr Glu 85 90 95

His Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met Gln Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Thr Leu Asn Cys Lys Asp Val Asn Ala Thr Asn Thr 130 135 140

Thr Asn Asp Ser Glu Gly Thr Met Glu Arg Gly Glu Ile Lys Asn Cys 145 150 155 160

Ser Phe Asn Ile Thr Thr Ser Ile Arg Asp Glu Val Gln Lys Glu Tyr 165 170 175

Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Thr 180 185 190

Ser Tyr Arg Leu Ile Ser Cys Asp Thr Ser Val Ile Thr Gin Ala Cys 195 200 205

Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala 210 215 220

Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Thr Phe Asn Gly Lys Gly 225 230 235 240

Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro 245 250 255

Val Val Ser Thr Gin Leu Leu Leu Asn Gly Ser Leu Ala Giu Giu Giu 260 265 270

Val Val Ile Arg Ser Asp Asn Phe Thr Asn Asn Ala Lys Thr Ile Ile 275 280 285

Val Gin Leu Lys Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn 290 295 300

Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr

The Sity Situ Ite Site Gly Asp Site Arg Sith Ata His Cys Ash Site Ser 325

Arg Ata Lys Trd Ash Asp Shr Leu Lys Gln Site Val Site Lys Leu Arg 345

Situ Gin Phe Glu Ash Lys Shr Ser Phe Ash His Ser Ser Gly Giy Asp Pro Glu Ite Val Het His Ser Phe Ash Cys Gly Gly Glu Phe Phe 370

Tyr Cys Ash Ser Thr Gin Leu Phe Ash Ser Thr Trd Ash Ash Ash Thr 385

Glu Gly Ser Ash Ash Thr Glu Gly Ash Thr Ite Thr Leu Pro Cys Arg 405

Ite Lys Gln Site Site Ash Met Val Gln Glu Val Gly Lys Ala Met Tyr 430

Ala Pro Pro Site Arg Gly Gin Site Arg Cys Ser Ser Ash Site Thr Gly 435

Leu Leu Leu Thr Arg Ash Gly Gly Site Ser Ser Ash Site Thr Gly 455

Tyr Lys Tyr Lys Val Val Lys Site Glu Pro Leu Gly Val Ala Pro Thr 485

Lys Ata Lys Arg Arg Arg Val Val Gln Arg Glu Lys

Lys Ata Lys Arg Arg Arg Val Val Gln Arg Glu Lys

Lys Ata Lys Arg Arg Arg Val Val Gln Arg Glu Lys

- (2) INFORMATION FOR SEQ ID NO:29:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 15 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: Linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
  - Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gin Arg Glu Lys Arg 1 10 15

5

## What is claimed is:

- A recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(W->X)</sub> point mutation, wherein X is an amino acid residue other than tryptophan.
- The recombinant nucleic acid molecule of claim 1,
   wherein X is a valine residue.
  - 3. The recombinant nucleic acid molecule of claim 1, wherein the nucleic acid molecule is a DNA molecule.
- 15 4. The recombinant nucleic acid molecule of claim 3, wherein the DNA molecule is a plasmid.
- 5. The recombinant nucleic acid molecule of claim 4, wherein the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
  - 6. The recombinant nucleic acid molecule of claim 1, wherein the C4 domain is an HIV-1<sub>IM</sub> gp120 envelope glycoprotein C4 domain.

- 7. The recombinant nucleic acid molecule of claim 6, wherein the mutant HIV-1 gp120 envelope glycoprotein is a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein.
- 30 8. The recombinant nucleic acid molecule of claim 1, wherein the C4 domain is an HIV-1 $_{\rm IR-FL}$  gp120 envelope glycoprotein C4 domain.
  - 9. The recombinant nucleic acid molecule of claim 8,

WO 94/22477 PCT/US94/03282

wherein the mutant HIV-1 gp120 envelope glycoprotein is a mutant HIV-1 $_{\rm JR-FL}$  gp120 envelope glycoprotein.

88

- 10. The mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of claim 1.
  - 11. A vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of claim 10, and an adjuvant.

10

12. A method of treating an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of claim 11, thereby treating the HIV-1-infected subject.

15

- 13. A vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of claim 10, and an adjuvant.
- 20 14. A method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of claim 13, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

25

- 15. A method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of claim 13, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.
- 16. A method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of

WO 94/22477

5

- HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of claim 13, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein.
- 10 17. The method of claim 16, wherein the subject is a human.
  - 18. The partially purified antibodies produced by the method of claim 16.
- 19. A pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of claim 18, and a pharmaceutically acceptable carrier.
- 20 20. A method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of claim 19 effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.
  - 21. A method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of claim 19 effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.
- 22. A composition which comprises a prophylactically effective amount of the partially purified antibodies of claim 18, and a pharmaceutically acceptable carrier.

5

- 23. A method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises administering to the HIV-1-exposed subject a dose of the composition of claim 22 effective to reduce the population of HIV-1 in the HIV-1-exposed subject, thereby reducing the likelihood of the subject's becoming infected with HIV-1.
- 10 24. The method of claim 23, wherein the subject is a medical practitioner.
  - 25. The method of claim 23, wherein the subject is a newborn infant.

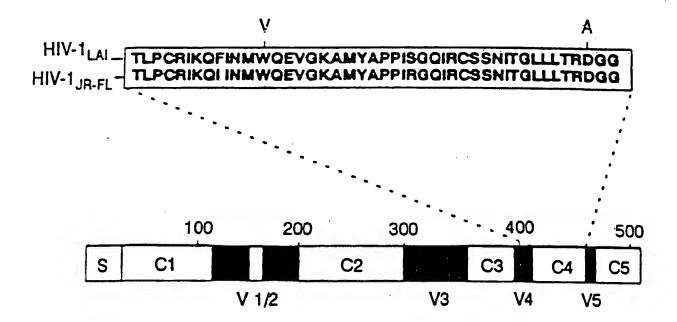
26. A method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a

becoming infected with HIV-1.

result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which

- comprises administering to the subject immediately prior to the incident a dose of the composition of claim 22 effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's
  - 27. The method of claim 26, wherein the subject is a medical practitioner.

## FIGURE 1



## FIGURE 2

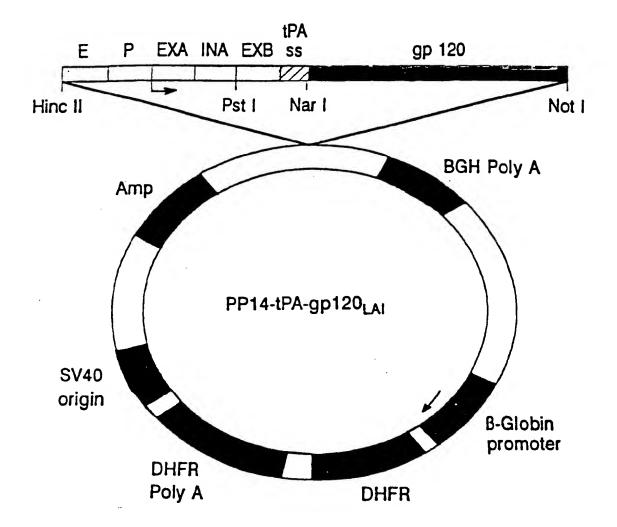


FIGURE 3A

FIGURE 30 FIGURE 3D FIGURE 3E FIGURE 3F

attagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactc »connatttccaactctccaccccattgacgtcaatgggagtttgttt	361
taaatggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgt	289
gtanactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacgg	217
aataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacg	145
gttccgcgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacgtc	73
ttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatgga	-

aagcagagetegtttagtgaaccGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAG ctecaegegaatetegggtaegtgtteeggaeatgggetetteteeggtageggeggageteeaeateegag cctgtcccatgcccatgcctccagcggctcatggtcgctcggcagctccttgctcctaacagtggaggccag tccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatat **ACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACGCGGATTCCCCCGTGCCAAGA**GTGA attggctatatgccaatactctgtccttcagagactgacacggactctgtattttacaggatggggtccca tttattatttacaaattcacatatacaacaacgccgtcccccgtgcccgcagtttttattaacatgcgggat acttaggcacaggacaatgcccaccaccagtgtgccgcacaaggccgtggcggtagggtatgtgtctga Cytaaytaccycctatagactctatagycacacccctttgyctcttatycatgctatactytttttgycttg ggccaacaccccgtcctagataggtgatggtatagcttagcctataggtgtggggttattgaccattattgac cactecectattggtgaegataetttecattaetaatecataaeatggeegetettgeeaeaaetatete ▼ Transcription Start Intron A **Exon A** 505 1009 649 577 793 865 721 937 1081 1153 1225

4/42

asstgageteggagattgggetegeaeegetgaegeatggaagatggaagaettaaggeageggeagaagaagatge 1297

FIGURE 3C

aggcagctgagttgttgtattctgtagagttggaggtaactcccgttgc@gtgctgttaacggtggagggca 1369

gtgtagtctgagcagtactcgttgctgccgcgcgccaccagacataakagctgacagactaacagactgt 1441

teettteeatgggtettte<del>etge</del>agTCACCGTCCTTGACACGATGGAATGAAGAGAGAGGGCTCTGCTGT tPA signal sequence Exon B PstI 1513

NarI 1585

agaacagaaaaattgtgggtcacagtctattatggggtacctgtgtgga\ggaagcaaccaccactctatt æ u × ш O ۵, Ø G **>** ட 4 O U H -1 .1 1657

×

H

.-

Signal cleavage

X H Z ы Δ **>** 4 × ~ Ω ဟ 1729 59

GACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGGAAAAATGACATGGTA Σ z H > ы a 1801

(RULE 26)

## FIGURE 3D

gaacagatgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtgtaaaattaaccccactc tgtgttagtttaaagtgcactgatttggggaatgctactaataccaatagtagtaataccaatag CTAGCAGÍAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTG ggggaaatgatgatgaaaaagaagagataaaaactgctctttcaatatggcacaagcataagataag GTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGAYAATGATACTACCAGCTATACG TATTGTGCCCCGGCTGGTTTTGCGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACA TTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACAT aa tgtcagcacagtacaatgtacacatggaattaggccagtagtatcaac'icaactgctgttgaatggcagt S G ပ Z × H H ഗ H ပ z H Ω ഥ H a ပ S S Z Z × \_. ~ . ഗ ۵. H 4 × Z ۲ S z H Z **,** H × × > م 0 Z S > ဟ Z H <u>ρ</u> H H a ပ م ပ Z H بعا Ω K K ပ Z Ω **¤** Z z 3 × .. H 0 × ~ ပ H H H -1 × v ဟ H က (L) × I X ٦ ٢ ₽ H ၒ > بعا × ഗ Ĺ ပ (a, ပ Δ M ၒ H 0 4 K × > M. Σ Σ z Ш .. ۵, ပ H Ш K I Ś S ы G 1873 2089 1945 2017 2305 155 179 2233 2377 131 2161 203 227 251

6/42

SUBSTITUTE SHEET (RIII F 26

FIGURE 3E

2449	MCCN	ATCI	GTA	<b>1 1 1 1 1 1 1 1 1 1</b>	ATT	AAT	TCT	3	AGAC	SS	25	<b>SAC</b>	<b>LTA</b>	3	KGA	<b>₩</b>	<b>AGT</b>	ATC	CGT	TC	CAG	<b>8</b> 0	SS.
	NOSVEINCTPPNNTPKSIRIOPG	ဟ	>	M	H	z	ပ	<b>(-</b>	œ	<u>م</u>	z	z	z	H	œ	×	ဟ	H	æ	н	0	<b>~</b>	ပ
	CCAGG	SAG	25	TTT	GIT	2	ATA	<b>₹</b> 95	3	ATAC	362	MI	ATG	AGA	SAA.	Š	ZAT	TGT	AAC	TT	T'U	40	<b>4</b>
323	PGRAFVTIGNMRQAHCNISRA	<b>64</b>	~	<b>[4</b>	>	۲	<b>H</b>	ပ	×	н	ပ	z	X	æ	ø	K	<b>=</b>	ပ	z	н	S	α	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \
2593	AAATG	SAN	ည်	ACT	TTA	X	3	ATA(	GCT	Ş	3	T X	୍ତ ଧୁ	<b>S</b> ₩S	3	TTT(	. 25 	MAT	MAT	₩.	ACA CA	ATA	ATC
347	K N N A T L K Q I A S K L R E Q F G N N K T I I	Z	<	H	H	×	0	H	4	ഗ	×	H	æ	(L)	0	Ĺ	G	z	z	×	۲	н	H
2665	TTTAA(	3	MTCC	5	39	99	SO	CCA	3	ATT(	STA	S	25	AGT	TTT	MAT	rgt(	SG.	9	***	TTT	TTC	TAC
371	FKOSSGGDPEIVTHSFNCGGEFFY	0	S	s	U	U	Ω	۵.	ы	н	>	<b>(-</b>	æ	တ	Ĺ	Z	ပ	ပ	U	ω	Ĺ	(a,	>-
2737	TGTAATTCAACACACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACA	TT	SS S	3	CTG	TIT	MAT	AGT,	ACT:	1551	rtti	MAT	AGT	ACT.	166	AGT,	ACT	GAA	999	TCA TCA	AAT	<b>₩</b>	ACT
395	Z U	S	H	0	H	(a <sub>4</sub>	Z	ဟ	H	3	[a,	z	ဟ	₽	3	လ	H	M	ပ	လ	z	z	<b>(-</b> -
2809	GAAGGAAGTGACACACACACTCCCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAA	<b>Z</b> AG:	IGAC	5	ATC	2	CIC	CCA	<b>1</b> 62	KG.	ATA	¥	3	TTT	ATA	<b>₹</b>	ATG.	<b>1</b> 66	CAG	SA	GTA	799	*
419	ш	S	Ω	H	<b>H</b>	H	<b></b>	۵۰	ပ	œ	<b>H</b>	×	a	[e.	H	z	Σ	3	ø	ല	>	<del>ن</del>	×
2881	GCAATGTATGCCCCTCCCATCAGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGA	GTA	IGCC	2001	ည်	ATC	AGC	<b>8</b> 8	CAN	ATT	AGA:	TGT.	ICA.	ICA	AAT.	ATT	ACA	999	CTG	CTA	TTA	AC.	AGA
443	X	<b>&gt;</b>	4	۵,	<u>α</u>	Ĥ	တ	ပ	0	н	œ	ပ	တ	S	Z	H	۲	ပ	H	J	<b>.</b>	H	œ.
2953	GATGGTGGTAATAACAACAATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA	<b>166</b> :	IAAT	[ <b>X</b> C	:AAC	MAT	99	<b>1</b> 00	GAG	ATC	ITC	AGA(	CCT	3GA(	3GA(	363	SAT	ATG.	AGG	SAC	<b>₹</b>	<b>1</b> 60	AGA
467	N N O O	G	Z	z	Z	z	ပ	ဟ	ш	H	Ĺ.	æ	۵	ပ	ပ	ၒ	۵	Σ	œ	۵	z	3	Œ

FIGURE 3F

~ × Ç H Ш × GTGGTGCAGAGAAAAATGAGCGCCCC NotI > ы æ . 0 > တ 3025 3097

FIGURE 4

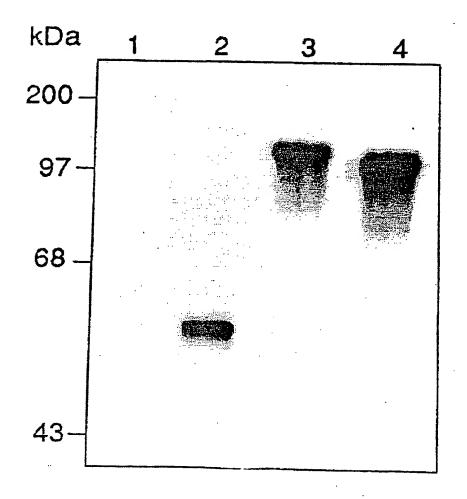
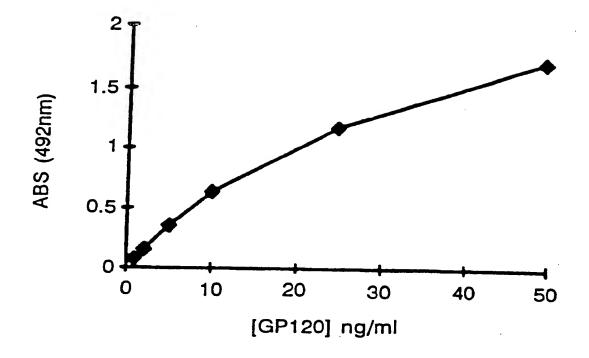


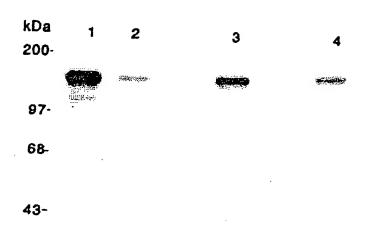
FIGURE 5A

Stable CHO	[gp120]
clone	(ng/ml)
5	6
6	14
9	123
10	4
12	18
13	18

FIGURE 5B



## FIGURE 6



29-

FIGURE 7A

JR-FL

FIGURE 7C FIGURE 7A FIGURE 7B

ATGGATGCAATGAAGAGA

≥ 4 Ω

Catgecegattcagaagagagagagaatagaaagtegegteacagtetatagg O ဟ ۵ 0 ш O U r r NarI U U O 27

₹ Signal ⋖ G Œ Œ 4 Œ ⋖

G

GTACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATAT S cleavage ⋖ U H 4 ы × ¥

GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCA Δ ρ, > U ~ = < \* > Z × L

Caagaagtagtattggaaatgtaacagaacattttaacatgtggaaaataacatggta E Z Ш H > Z ш > O 259 87 gaacagatgcaggagatataatcagtttatgggatcaaagcctaaagccatgtgtaaaa × H S 0 Ω \* H S H Ω W X, 0 319

**TTAACCCCACTCTGTGTTACTTTAATTGCAAGGATGTGAATGCTACTAATACCACTAAT** Z < Z > Ω × ပ Z \_1 > ပ 379 127 GATAGCGAGGGAACGATGGAGAGGAGAAATAAAAAACTGCTCTTTCAATATATCACCACA S O W G X

439

SUBSTITUTE SHEET (RULE 26)

139

47

199

# FIGURE 7B

**AGCATAAGAGATGAGGTGCAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA ATAGATAATAATAACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAG** GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT gcgattctaaagtgtaatgataaggtcaatggaaaggaccatgtaaaaatgtcagc **AGAAAAAGTATACATATAGGACCAGGGAGAGCATTTTATACTACAGGAGAAATAATAGGA AGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTTCACGAAGAATGCTAAAACC ATAATAGTACAGCTGAAAGAA**TCTGTAGAAATTAATTGTACAAGACCCAACAACAATACA gatataagacaaattgtaacattagtagagcaaaatggaatgactttaaaacag z M S ပ Ö 0 × H Ω G F٠ ပ တ ပ × z **>** ပ S Ω > z Ĺ .-> H Z S **-**4 ₄۵ œ, م æ .. u S œ M œ > ပ بم H H × > တ W ۵, z Ω 9 = ပ > ш 0 S ပ S Z × W H H ပ -1 **=** Z L ပ × Z × W ... ہم 0 z S 499 559 619 619 739 167 187 207 227 247 799 267 859 287 919 979 307 327

14/42

C	)
1	•
5	
Ľ,	,

CCTGGAGGAGATATGAGGACAATTGGAGAAGTGAATTATATAAAATAAAGTAGTA P G G G D M R D N W R S E L Y K Y K V V AAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAAGAAGAGAGTGGTGCAAAGAAA K I E P L G V A P T K A K R R V V Q R E Noti
GAGGA G G AACCA NotI

FIGURE 8A

FIGURE 8A FIGURE 8C FIGURE 8B

LAI AV3

ATGGATGCAATGAAGAGAGGCCTCTGCTGTGTGTG U H

ctgctgtgtggagcagtcttcgtttcgcccagccaggaaatccatgcccgattcagaagaggggccagaaca U ø ш O

Signal cleavage H Ш × > ۵. > დ ⊁ **>** > 109

TCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCC <u>بم</u> > ပ 4 I H ~ 3 > z > Ш Ω **>** 4 × 4 က 181 61

aacccacagaagtagtattggtaaatgtgacagaaaattttaacatgtggaaaaatgacatggtagaacag Z × Σ Z Ы H > Z H > Ш 0 253 85

atgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtgtaaaattaaccccactctgtgtt × -1 S 0 325 109 agtitaaagtgcactgatttggggaatgctactaataccaatagtagtaataccaatagtagtagcgggaa S

S G 1 ပ 397 133

FIGURE 8B

**ATGATGATGGAGAAAGGAGAGATAAAAAC**TGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAG GCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTC GAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTGAACCAA **aaagaatatgcattttttataaa**cttgatataataccaatagataatgatactaccagctatacgttgaca <u> Agttgtaacacctcagtcattacacaggcctgtccaaaggtatcctttgagccaattcccatacattattgt</u> tctgtagalattaattgtacaggtgctggacattgtaacattagtaggggcalaatggaatgccactttaaaa ပ S × ပ G ပ 4 > æ Z Z H -1 H ပ 3 H H တ H H م × ⊢ Δ G -1 H 回 4 ഗ 0 × Z æ Z ~ H ₽ Δ ഗ Ĺų ഗ z ဟ Z H H H Δ Ĺ > ۵. × ¥ z > H ഗ H z ۵, U ပ z م بعا ပ H 岡 æ ပ ~ Z Ω Z G × 4 ,, × a H ပ H × လ H U H **= >** æ ы H ~ H H G (zą U ပ ¥ > (a, S 4 ပ > > **>** 4 W Σ W W S × Ы 829 469 613 685 229 253 277 901 301 757 157 205 181 541

SHEET (RIII F 26)

SUBSTITUTE

FIGURE 8C

#### 18/42

CAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGG GACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACACTGTTT Ω ဟ S ഗ 0 G × z ) ---EH Ĺų z z ഗ × W G z G z ပ Ы ပ ပ H لعا z S 0 3 Ĺ M တ H œ S × H H Z × > 3 u H S 973 325 1045 349 1117 373

**CTCCCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAAGCAATGTATGCCCCTCCCAT**C agcggacaaattagatgttcatcaaatattacaggctgctattaacaagagatggtggtaataacaacaa K Σ ~ × G > M 0 3 Σ z Н Ĺ., 0 × Н æ ပ 1189 397 1261

**GGGTCCGAGATCTTCAGACCTGGAGGAGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATA**TAAA ы ဟ œ \* z Ω ø, Σ Δ U O G Δ, **~** بعا H 'n S G 1333 445

G

Ö

۵ ۳

H

H

H

z

S

ပ

œ

H

0

ග

421

Ш æ × ы > 1405

NotI

1477 GOGGCGC

ATGGATGCAATGAAGAGA GGGCTCTGCTGTGTGCTGCTGTGTGTGAGCAGTCTTCGTTTCGCCCAGCAGGAAATC CATGCCCGATTCAGAAGAGGCGGCAGAGTAGAAAGTTGTGGGTCACAGTCTATTATGGG GTACCTGTGTGGAAAGAAGCAACCACCACTČTATTTTGTGČATCAGATGCTAAAGCATAT GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACCAGACCCCAACCCA Caagaagtagtattggaaatgtaacagaacattttaacatgggaaaaataacatggta GANCAGATGCAGGAGGATATAATCAGTTTATGGGATCAAAGCCTALAGCCATGTGTAAAA **TTAACCCCACTCTGTGTTACTTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT** ш O FIGURE 9C FIGURE 9A FIGURE 9B 4 Ω **80** × ဟ cleavage Þ S U Z O Ω 0 **FIGURE 9A** Ш ⋖ × G U ບ S NarI ⋖ > O Z H ø Ш Z W Ω ×  $\Xi$ Σ < JR-FL AV3 19 139 199 67 79 259 87 47 319 379 127

# **GURE 9B**

39	GATAGCGAGGGAACGAGAGAGAGAATAAAAACTGCTCTTTCAATATCACCACA D S E G T M E R G E I K N C S F N I T T
99	AGCATAAGAGATGAGGAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA S I R D E V Q K E Y A L F Y K L D V V P
87	ATAGATAATAATACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAGIS DE DE NEN NET SEN LE SEN DE SEN LE DE SEN LE DE
619 207	GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT A C P K I S F E P I P I H Y C A P A G F
679 227	GCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC A I L K C N D K T F N G K G P C K N V S
39	ACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGCTAAATGGC T V Q C T H G I R P V V S T Q L L L N G
99	AGTCTAGCAGAAGAAGAGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC S L A E E E V V I R S D N F T N N A K T
59 78	ATAATAGTACAGCTGAAAGAATCTGTAGAAATTAATTGTACAGGTGCTGGTGGACATTGTAAC I I V Q L K E S V E I N C T G Å G B C N
919 307	ATTAGTAGAGCAAAATGGAATGACACTTTAAAACAGATAGTTATAALATTAAGAGAACAA I S R A K W N D T L K Q I V I K L R E Q

20/42

_	
Ç	ر
C	n
Ц	ل
	ζ
=	)
C	5
ū	_

							:
TTTGAGAATAAAAAATAGTCTTTAATCACTCCTCAGGAGGGGACCCAGAAATTGTAATG FENKTIVFNHSSGGDPEIVM	S	A P	P P	SAT	rattaatgagaaccgagatcttcagacctggaggagagatatgagggac Ineengte Ifrpgg	AATIGGAGAAGTGAATTATAAATATAAAGTAGTAAAATTGAACCATTAGGAGTAGCA N W R S E L Y K Y K V V K I E P L G V A	
STA	Z Z	H 1	) } <b>~</b>	R GA	0 0 <b>%</b>	GT3	
I	TIA	CA F	ATG	Ş U ⊢	A G	ဗို့ ပ	
E	ıGı	2	161	<b>\( \)</b>	ATA C	ĬŢ.	••
CAG P	Ş	¥	\$	AT	AG.	CA a.	g
	2	TA T	96	GCI	Q G	<b>∑</b> 3	<b>8</b> 81
<b>1</b>	Ծ ₽	§ z	3 ×	SC 1	g o	Ö.	<b>-8</b>
<b>8</b> 66	ည်း	ကြွ	ပြွ ပ	ပ္ပံ့ပ	ပြွဲ ပ	AA1	TGA -
9	Z z	<b>₹</b> ω	GT2 >	A CA	ည်ရ	\$ ×	<b>3</b> ×
TC)	161 C	ACT	SAA	ATT	N S	VGT.	AAGGCAAAGAGAGAGTGGTGCAAAGAGAAAATG K A K R R V V Q R E K -
TCC S	Y Y	<b>5</b> 2	80	Z Z	ii w	CT7	<b>1 1 1 1 1 1 1 1 1 1</b>
CAC	ii ii	. 4 z	ပ္ပံ ဖုံ	ည္သ	E H	₹×	<b>\$</b> a
Z Z	E a	ညီ	75. K	CAT	AGA E	TAT X	ပ္ညည
F	¥ 3	GGT	A N N	STT	S S S	Ž×	9
707	SAG G	<b>5</b> 3	<b>3</b>	SAT	Š	X X	NG.
TAG I	<b>Y</b> 5 (5	150	TA	TAC	Se constant	IAI	<b>X</b>
\$	70	TAC	A I	MAT	3 z	AT E	34G R
<b>\$</b> [	ori O	X z	ర్జ్ఞ	g o	Ş W	15 to 1	₹×
£ ×	Ž z	Ž z	₹ ×	ည်ပ	ŽZ	<b>3</b>	र्धुं ∢
Z Z	E G	Ž z	ATA I	AGA R	ATT	GAG	\ <b>\S</b> ×
E E	AGT S	100 100	<b>5</b> ~	CATC	GGT	Ğ. ¥	Q F
F	CACAGITITAATIGIGGAGGAGAATITITCIACIGIAATICAACACAACA	ACTTGGAATAATACTGAAGGGTCAAATAACACTGAAGGAAATACTATCACACTCCCA T W N N N T E G S N N T E G N T I T L P	TGCAGAATAAAACAATTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCT C R I K Q I I N M W Q E V G K A M Y A P	CCCATCAGAGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGAT PIRGOIRCSSNITG LLLTRD	GGTGG1 G G	\{\frac{1}{2}} \text{z}	NOTI CCCACCAAGGCAAAGAGAGAGTGGTGCAAAGAGAAAAATGAGCGGCCGC P T K A K R R V V Q R E K -
	•			J	J		O
327	039 347	99	159 387	61	9 C	39	99 87
ש אי	1039 347	1099 367	38	1219	1279	133 44	139

FIGURE 10A

LAI AV3-CD4

FIGURE 108 FIGURE 10C FIGURE 10A

**ATGGATGCAATGAA**GAGAGGGCTCTGCTGTGTGCTG ပ

ctgctgtgtggagcagtcttcgtttcgcccagccaggaaatccatgcccgattcagaagaggggccagaaca NarI O <u>م</u> . د 4 ш O م ဟ Þ ட Þ **∀ U** ပ H

Signal cleavage 臼 × ۵, G W 109

TCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCC ပ 4 ¥ H Z Z = > Ш H Ω × 4 × 4 Ω ഗ 181 61

**AACCCACAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGGAAAAATGACATGGTAGAA**CAG Ę Ω Z 3 E Z <u>L</u> Z Ы > z > -1 > > M Ú 253 85

atgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtaaattaaccccactctgtgtt H **×** > U ۵, × .. S 0 Ω 3 -1 S Н ۵ ú × 325 109

<u>AGTITAAAGIGCACTGATITGGGGAATGCTACTAATACCAATAGTAGTAATACCAATAGTAGTAGTAGCGGGGAA</u> G လ Z Z တ Z Z z G 397 133

TUTE SHEET (RULE 26) FIGURE 10B

**CAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCA**GGAGGG

**~** 

U

U

301

G

ပ

ဟ

0

×

H

H

×

Z

G

0

Ш

∝

H

×

973 325

# 23/42

atgatgatgagaaaagagagataaaaaactgctctttcaatatcagcacaagcataagaggtaaggtgcag **aaagaatatgcatttttttaaacttgatataataccaatagataatactaccagctatacg**ttgaca <u> Agttgtaacacctcagtcattacacaggcctgtccaaaggtatcctttgagccaattcccatacattattgt</u> GCCCCGGCTGGTTTTGCGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTC **AGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTGAATGGCAGTCT**AGCA GAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTGAACCAA tctgtagaaattaat<del>agtacaggtgctggacattgt</del>aacattagtaggcaaatggaatgccactttaaaa 0 ပ G ىم z > ၒ H H <u>م</u> . . ഗ H ပ Ω H S Z [sa Z O × Ω တ [a, H 4 တ H H > z ဟ Z × × م > z ۵. H z > U H ပ z م 4 ۵ ပ æ z z \_ × o × H ບ 4 J × H S U ы = œ ~ > H G ပ S Ĺ. × G (L) H 0 Σ z Σ ပ Σ 829 469 157 613 205 685 229 541 253 181 757 277 901

FIGURE 10C

GACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTT aatagtacttggtttaatagtacttggagtactgaagggtcaaataacac†gaaggaagtgacacaatcaca **CTCCCATGCAGAATAAAACAATTTATAAACATGGTGC**AGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC agcggacaaattagatgttcatcaaatattacaggctgctattaacaagagatggtggtaataacaac Z 4 Ω Z ഗ ဟ **>** ပ G Σ G M æ Δ × æ G H z > z H H M (L) ဟ H ပ 0 G ပ ပ > (L) ပ Σ H H z H ഗ Z, Ĺ 3 z S S Ĺa, I 0 S S × H ပ Z æ H æ O ပ ഗ 1045 349 1117 373 1189 397 1261 421

**GTAGTAAAAA**TTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAATGA > > × H ۵, 4 > G H O. W H × > 1405 469

**GGGTCCGAGATCTTCAGA**CCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAA

Ы

S

æ

3

Z

Ω

œ

X

Ω

G

ပ

ပ

Δ,

æ

Ĺ,

H

ᆸ

G

445

1333

NotI

2000000 1447

FIGURE 11A

JR-FL AV3-CD4<sup>-</sup>

ATGGATGCAATGAAGAGA  M D A M K R	GGGCTCTGCTGTGTGTGTGGAGCAGTCTTCGTTTCGCCCAGCCAG	CATGCCCGATTCAGAAGAGCGCCAGAGTAGAAAAGTTGTGGGTCACAGTCTATTATGGG  H A R F R R G A R V E K L W V T V Y G A Signal alemana	GTACCTGTGTGGAAAGAAGCAACCACCACTGTTTTTGTGCATCAGATGCTAAAGCATAT  V P V W K E A T T T L F C A S D A K A Y	GATACAGAGGTACATAATGTTTGGGCCACACACACTGTGTACCCACAGACCCCAACCCA D T E V H N V W A T H A C V P T D P N P	CAAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGTA Q E V V L E N V T E H F N M W K N N M V	GAACAGATGCAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAA E Q M Q E D I I S L W D Q S L K P C V K	TTAACCCCACTCTGTGTTACTTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT  L T P L C V T L N C K D V N A T N T T N
44	19	79 27	39	99	59 87	19	79

	1	
1	_	_
1	1	1
[	1	_
7	•	) 3
ì	_	_
_		

	GATAGCGAGGGAACGATGGAGAGAGGAGAAATAAAAAACTGCTTTTCAATATACAA
147	DSEGTMERGEIKNCSFNITT
	AGCATAAGAGATGAGGAAAGAATATGCTCTTTTTTATAAAAAAGAATATA
167	S I R D E V Q K E Y A L F Y K L D V V P
S	ATAGATAATAATACCAGCTATAGGTTGATAAGTTTGTGAGAGGGGGGGG
187	I D N N T S Y R L I S C D T S V I T Q
~	GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCATACATTATTATTGTGCCCATACATTCCTTTCCTTTCTTCTTCTTCTTCTTCTTCT
207	ACPKISFEPIPIHYCAPAGF
619	GCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGACA
227	AILKCNDKTFNGKGPCKNVS
C	ACAG!'ACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTAAATGGC
247	TVQCTHGIRPVVSTQLLLNG
199	AGICTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTTCACGAAQAATGCTAAAACC
267	SLAEEEVVIRSDNFTNNAKT
	ATAATAGTACAGCTGAAAGAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC
287	IIVOLKESVEINCTGAGACN
919	ATTAGTAGAGCAAAATGGAATGACACTTTAAAACAGATAGTTATAAAAITTAAGAGAACAA
307	ISRAKWND TLKOIVIKLREO

# FIGURE 11C

979 327	TITGAGAATAAAACAATAGICITIAAICACICCICAGAGGGGGACCCAGAAATIGIAAIG	X X	\$	IAG I	77.	TT	Ĭ.	CAC	ភ្ជិ	J. C	3	S c	SGA	ບູ	AGA	\$	TTG	₹:	TG	
	<b>.</b>	•			•	4	:	=	9	0	9	9	<b>a</b>	۔ بد				>	Σ	
1039	CACAGITITA	ATTG	ii ii	3AG	3AG	AAT	TT	TC	TAC	IGI	AAT	Ç		5	ACT.	F E	Į.	4 E	£	
347	H S F N C G G E F F Y C N S T Q L F N S	z		,,	(5)	ш	(Le,	£a,	<b>&gt;</b>	ပ	z	လ	H				-		, , ,	
1099	ACTIGGAATA	ATA	TAC	75	S Z	361	3	7	ZY A	ACT	3	Ş	, A A	<u>.</u> A L	T & T	4	ز	į	ć	
367	TWNNTEGSNATEGNAILP	z		-	(-)	(2)	S	Z	Z	<b>E</b> →	ш	U	Z	) <b>⊱</b> -	H	<b>5</b>		֚֚֡֝֝֝֞֜֜֜֝֜֜֝֜֜֜֝֓֓֓֓֜֜֜֜֜֜֜֜֜֜֜֜֜֜֜֜֜֜֜	ج د د	
1159	TGCAGAATAA	<b>S</b> C	\{\{\}	TA	3	S S S	5 S	Š	Q X	3	GTA	ij	3	ပ္တ	Z.	GIS	) J	S C C	E C	
387	CRIKOIINM VOEVGK MMY AP	<b>∞</b>				7	Σ	>	0	M	>	G	×	~	X			<b>.</b>	, A,	
1219	CCCATCAGAGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGAT	SACA	Z	TA(	SAT	STT	2	3	ZI S	ATT	AC A	g	); (1)	S	ATT	₹	Ž	SAG	AT	
407	P 1 8 (	()	-		~ ·	•	ဟ	<b>တ</b> .	z	<b>H</b>	H	ပ	-1	-	H	-	-	œ	Ω	
1279	GGTGGTATTAATGAGAATGGGACCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGAC	ATGA	3	ST.	SSA	ပ္ပ	AG.	T C	T.	AGA	SCI	ÿ	199	3	AGA	TAT	GA	ပ္ပ	<b>A</b> C	
427	0 0 D	M	Z	_		<u>.</u>	ш	H	(4,	œ	۵,	ပ	<u>ල</u>	<b>5</b>	Ω	2,	_	~	Ω	
1339	AA'ITGGAGAA	AGTG	AT	TAI	AT	3	TAT	₹	(GT	AGT.	₹	X X	TG	Š	2 F	TAG	CAC	STA	SCA	
447	N W R S E L Y K Y K V K I E P L G V A	S	<b>ы</b>	H	<b>&gt;</b>	×	<b>&gt;</b>	×	>	>	×	H		-	۵.	H	ပ	>	4	
1399	Not I	.XXX	3 <b>A</b> G	<b>₹</b>	AG.	ပ္ပ	ပ္ပ	3	SA S	₹	₹	TGA	≃ ႘	Not I	S					٠.
487	PTKAKRKVVOREK	×	8					a	<b>~</b>	ш	×		) }	j	}					

FIGURE 12A

LAI CD4

FIGURE 12C FIGURE 12A FIGURE 12B

**ATGGATGCAATGAAGAGAGGGCTCTGCTGT** ပ tPA signal sequence 1 O

GTGCTGCTGCTGTGTGGAGCAGTCTTCGCCCCAGCCAGGAAATCCATGCCCGATTCAGAAGAGGCGCC Nari **K** ⋖ I Н ш O > ⋖ O U H

H K Ш × 3 > ۵, > G >-× > 3 .-× Ы 109 37

Signal cleavage 181

TGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCTTGTGTACCCACA I H 4 3 > Z = > Ы H Ω × K × **A** S 61

GACCCCAACCCACAAGAAGIAGIAITGGIAAATGTGACAGAAAATTTTAACATGGGAAAAATGACATGGIA z Z 回 > Z > > Ш 0 253 85 GAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTC ,, × > ပ × H S 0 0 3 H S 0 Ы E 325 109

TGTGTTAGTTTAAAGTGCACTGATTTGGGGAATGCTACTAATACCAATAG1'AGTAATACCAATAGTAG'TAGC S ഗ S Z Z ဟ S Z Z Z G H 0 H ပ × ,, ഗ 397 133

(RULE

G

0

æ

Н

×

œ

H

z

z

Z

م

æ

H

ပ

z

W

301

**CCAGGGAGAGCATTTGTTACAATAGGAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAG**KGCA

4

a

æ

Σ

z

G

¥

G

H

H

4

œ

G

325

973

# FIGURE 12B

TATTGTGCCCCGGCTGGTTTTGCGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACA CTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACA(%CTG **GGGGAAATGATGATGGAGAAAGGAGAGATAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAG GTGCAGAAAGAATATGCATTTTTTTTATAA**CTTGATATAATACCAATAGATAATGATACTACCAGCTATACG **AATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTGAACTGCTGTTGAATGGCAGT** TTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACAT aaccaatctgtagaaattaattgtacaagacccaacaatacaagaaaaaagtatccgtatccagagggga S z H ပ H .. H H ഗ ပ -1 H Ĺ Ω 0 × ഗ H 4 Ω ဟ H H z S z H > Δ × > × Ĺų Ω, H > ഗ H ۵, Z ပ ۵, Ç., Н ပ z z 4 ပ œ Z, Δ × H 0 × H 4 H H × ၒ S H ш × æ H > H H G 4 H Ĺ > ပ > S Ĺ. × Ĺų ၒ > Ö ~ u > 4 Ы E **>** Z I ۵, M L ပ 4 S S M × 4 0 ပ 829 685 229 277 469 613 205 253 901 157 757 541 181

29/42

FIGURE 12C

**AAATGGAATGCCACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAAT** AATC × z Z ပ Ĺ O W æ, L × ဟ 4 H 0 × \_\_ H ~ 1045

**TTTAAGCAATCCTCAGGAGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTT**TCTAC G O ပ z (a, ഗ I H > н Ы ۵, Ω ပ G ဟ ഗ 1117 373

TGTAATTCAACACAACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACT 1189 397

M S 3 H S Z [a, 3 H S Z Ĺ. H 0 H S

GCAATGTATGCCCCTCCCATCAGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGA GAAGGAAGTGACACACTCCCATGCAGAATAAAACAATTAAAACATGGTGAGGAAGTAGGA%AA ပ Σ z Ĺ 0 æ ပ <u>α</u>, .1 H H H Δ လ ග 1333 1261 421

GATGGTGGTAATAACAACAATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA Σ Ω ပ ပ ၒ Δ, œ بعكا ы Ņ G z z Z Z G G 1405 469

.-1

ပ

H

H

z

ဟ

S

ပ

æ

0

<u>ග</u>

ဟ

H

۵,

445

H 4 > ပ H Ш × × × × 1477 493

· NotI

**GTGGTGCAGAGAAAAATGAGGGGCCGC** 1549

0 517

L L Nari

ບ

FIGURE 13A

FIGURE 13A FIGURE 13B

FIGURE 13D		ATGGATGCAATGAAGAGA M D A M K R
	JR-FL CD4"	1 1

R F R R G A R V E K L W V T V Y Y G	TAT	V W K E A T T T L F C A S D A K A Y	Ş	<u>α</u>	CAAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGTA	>
<b>&gt;</b> +	Ş	4	Š	z	AT D	X,
<b>&gt;</b>	Ž	×	g	۵,	SC	z
>	N Z	~	NC ST	۵	Z Z	z
H	ATC	Ω	CAC	H	₹	×
>	Š	S	Ç	Δ.	3	3
<b>z</b> 6	35	~	TAC	>	.TGT	I
	GTO	ပ	GTG	ပ	5	z
× [	TI	Ĺu,	$\Sigma$	~	TT	Ĺu,
E E	TA	1	:ATG	×	XT.	<b>=</b>
ک <u>در</u>		H	3	H	3	ы
<b>~</b> <		H	$\frac{3}{2}$	~	3	H
⋖	Ç	H	8	*	Ki	>
U	3	4	TI	>	JATO	z
K	SAAC	Œ	\$TE	z	3	凶
ø	Ž	×	AT	×	T S	H
_	<u> </u>	3	STAC	>	ITA	>
<b>a</b>	STG	>	340	(L)	JTX(	>
	Ç	<u>α</u> ,	Š	H	X	ω
∢ H	GTACCTO	Þ	GNT	DIEVHNVWATHACVPTDPNP	25	<b>О</b>
27	6	7	9	2	ف	7
. `	14.	•	<b>a</b>	w	10	-

FIGURE 13E

<b>\$</b> ×	AAT N	A F	P 6	50	FIT
GAACAGATGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAA E Q M Q E D I I S L W D Q S L K P C V K	ACCCCACTCTGTGTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT T P L C V T L N C K D V N A T N T T N	AGCGAGGGAACGAGAGAGAGAATAAAAAACTGCTCTTTCAATATCACCACA S E G T M E R G E I K N C S F N I T T	ATAAGAGATGAGGAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA I R D E V Q K E Y A L F Y K L D V V P	ATAGATAATAATAATAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAGIID N N N T S Y R L I S C D T S V I T O	IGICCAAAGATAICCTITIGAGCCAATICCCATACATTATIGIGCCCCGGCIGGITIT C P K I S F E P I P I H Y C A P A G F
CC	ACC	ATC	GTA >	ATT	SCT A
స్ట్ చ	N	N	GAT	GTC V	ည
<b>X</b> ×	ACT T		- Co	ည်း	ညည
E 1	<b>1</b> 3€1	S	<b>₹</b> ×	<b>2</b>	TGT C
ည် နှ	ZAZ Z	ည် က	TAT	GAC D	TAT
ဥၟႄၜ	1610 V	3 2	rtti F	ក្ ក	EAT H
<b>19</b>	<b>3</b> 50	}×	- E	<b>V</b> 63	I
ATQ SE	¥ ×	AT.	ည် 🗸	EAT!	ည်
TTT T	រីក ភ	ξξ E	ATA Y	STIC	2 -
S S	Ž	<b>1</b> 66	AGA E	rage R	ည္တေ
<b>Ž</b> 1	ITT L	SAG!	AS ×	Y X	IGA E
IAI. I	TAC	GGA	ğo	S S	TT.
S C	TGT V	GAT	<b>661</b>	13C 1	ATC. S
Α55 Θ	ည် ဂ	Z T	TGA	A Z	GAT
<b>5</b> 0	ÄČĪ	ပ္ပ ပ	AS a	AAT	<b>₹</b> ×
X	ပ္ပ	SS	AS A	TAAT	ည် မ
हु <b>व</b>	Z F	TAG	S	AS a	ည်ပ
g w	TTV	GAT	AGC.	AT	GCC3 ▼
107	379 127	439	499	559 187	619
กลี	m H	4 4	<b>4</b> A	80 H	9 6

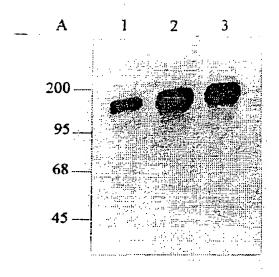
FIGURE 13(

လွှင	ပ္ပ ပ	T CC	J.	ဗွ် ဗ	80
μς > Ω	AAT N	<b>≩</b> ×	atagtacagctgaaagaatctgtagaattaattgtacaagacgcaacaacaataca I V Q L K E S V E I N C T R F N N N T	ramagtatacatataggaccagggagagagggagagaggaggagaataatagga Ksihig pgrafyttatactaggagaaataatagga	ATAAGACAAGCACATTGTAACATTAGTAGACAAAATGGAATGACACTTTAAAACAG I R Q A H C N I S R A K W N D T L K Q
Z Z	CT.	3CT.	<b>2 2</b>	ATA	TA 1
<b>₹</b> ×	CTG L	Z Z	<b>2</b> z	Ag u	ACT T
161 C	<u>က်</u> ၁	- <b>X</b> %-	ဂ္ဂ ဂ္ဂ	ပ္လို့ပ	ည်ရ
ပ္ပံ 🏎	<b>₹</b> ∘	ACG	AGA R	ACA T	MA
<b>လ</b> ပ	ACT T	TIC	H H	T T	17GG ¥
<b>3</b> ×	NTC. S	ZAT N	ည်း	rtati X	}×
ပြီ ပ	V V	O C	Z Z	ATT7 F	)   V
ž z	AGT.	ATCI	ATT I	IJ ✓	IAG/ R
STT	ည္တြေ	TAG. R	AGE :	GAG R	TAG
Ŭ B	TAG R	<b>*</b>	1GT.	<b>გ</b> ვ	S. I
Ž ×	₩.	AGT	ATC	ACC	A Z
40 T		199 V	AS 3	98	TTG
Z H	S =	AGA	3	TAT I	S. T.
	TAC	15 a	200	2 -	\(\frac{1}{2}\)
	215 0	36.	5	STA	25.
	5	36.		<b>3</b>	
GCGNTTCTARAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC A i l k c n d k t f n g k g p c k n v s	ACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGCTAAATGGC T V Q C T H G I R P V V S T Q L L L N G	AGTCTAGCAGAAGAGAGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC S L A E E E V V I R S D N F T N N A K T	ATAA	AGEAN R	GATA
<b>5</b>	¥.	<b>¥</b> "	*	<b>Z</b> –	<b>3</b> –
619 227	739	799 267	859 287	919	979

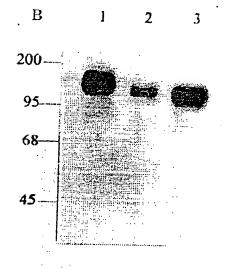
"			
		2	
L			l
ב		۰	
5		3	
Ĺ		-	

1039	ATA	GTT	ATA	*	TTA	AGA	₹ S	<b>₹</b>	TII	GAG	MAT	A	ACA	ATA(	i) L	FFF	FAA	ָר ה נ	j.	į
347	H	<b>&gt;</b>	н	×	H	œ	ы	0	Ĺ	IVIKEREFRATIVENHSS	z	×	H	н	>	1 1 - faz 1	z	S =	် က	S
1099	GGA	999	GAC	${f 5}$	\$5	ATT	GTA	ATG	Ş	AGT	TII	AAT	TGT	<b>4</b> 55		. A	4	E	Ė	į
367	G	ပ	Ω	Ω.	ш	н	>	Σ	×	GGDPEIVMHSFNCGGEFFYC	[E <sub>4</sub>	z	ပ	5	၌ ပ	ξω.	į (m	, 1 1	X X	5 U
1159	AAT	TCA	ACA	3	CTG	TII	AAT	AGT	ACT	166	<b>MAT</b>	MI	ZY.	ACT	GAA	) ()	Ţ	AAT	, A A	ز
387	Z	ഗ	H	O	H	£4,	z	S	H	NSTOLFNSTERNATEGSNATINAT	z	Z	z	H	M	<b>9</b>	လ	Z	Z	
1219	GAS	<b>1</b> 55	AAT	ACT	ATC	ACA	CIC	S	35	AGA	ATA	Z	3	ATT	ATA	A C	ATG	616	Ş	Z S
407	េ	Ġ	z	H	H	H	н	۵	ပ	EGNTITLPCRIKQIINMVOE	H	×	o	н	н	2	Σ	<b>&gt;</b>	a	四
1279	GIA	<b>199</b>	Ž	ర్ట	ATG	TAT	ပ္ပ	S	ပ္ပ	ATC	AGA(	SGA G	3	ATT	AGA:	IGI	ŢĊ	, T	3	CAT
427	>	<u>ල</u>	×	<b>«</b>	X	<b>&gt;</b> 4·	K	ىم	۵	V G K A M Y A P P I R G Q I R C S S N I	œ	ပ	a	н	œ	ပ	ဟ	S	Z	H
1339	ACA	99	OHU OHU	CIA	TTA	NO.	AGA	GAT	GT	<b>ACAGGCTGCTATTAACAAGAGATGGTGGTATTAATGAGAATGGGACCGAGATCTTCAGA</b>	ATT	MI	GAG	Z Z	999	ည္ခ	GAG	ATC	H	<b>SA</b>
447	۲	ပ	1	H	ы	H	α,	Ω	<b>o</b> .	T G L L T R D G G I N E N G T E I F	H	Z	M	z	ၒ	Ħ	M	H	(m,	æ
1399	CCT	3GA(	3GA	GA GGA	GAT	ATG	<b>1</b> 00	SAC	<b>₹</b>	CCTGGAGGAGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATAA	AGA.	AGT	SAA:	ITA.	IAT	X	TAT	3	GT	NGT.
467	۵.	ၒ	ပ	ပ	۵	Σ	<b>~</b>	Ω	Z	GGDMRDNRRSELYKKV	æ	S	山	H	<b>&gt;</b>	×	<b>&gt;</b>	×	>	>
1459	AAA	ATT(	3 <b>XX</b> (	S.	ITA(	3GA(	3TA(	SS.	S	<b>AAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAGAG</b>	४४	3	<b>M</b> G	3	AGA(	ST G	GTG	3	Z Z	3
48	×	<b>H</b>	m :	ا بم	H	ပ	>	4	م	EPLGVAPTKAKRWVOR	×	~	×	æ	<b>~</b>	>	>	a	æ	ы
1519	NotI AAATGAGCGGCCGC	GAG	<u>දු පූ</u>	Section 2015	ĸ															
507	×																			

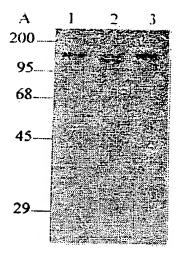
### FIGURE 14A



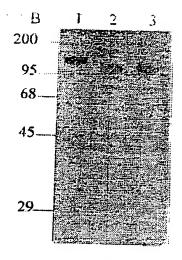
### FIGURE 14B

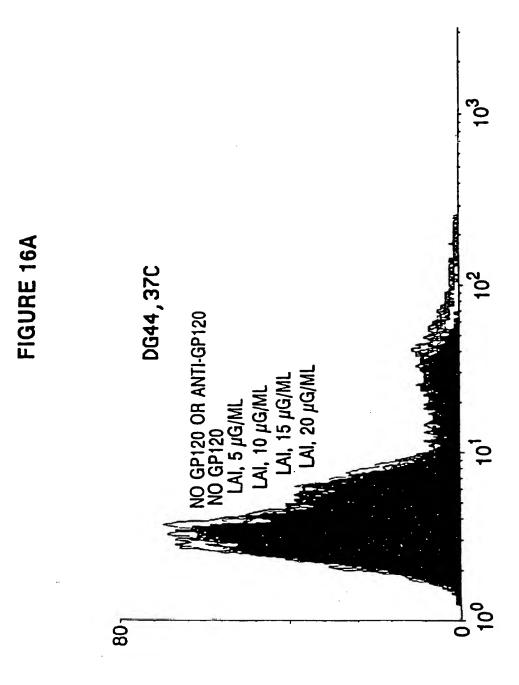


#### FIGURE 15A

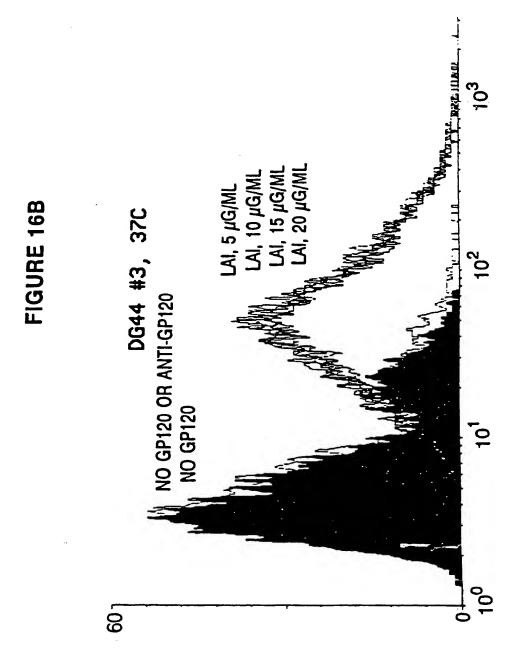


#### FIGURE 15B



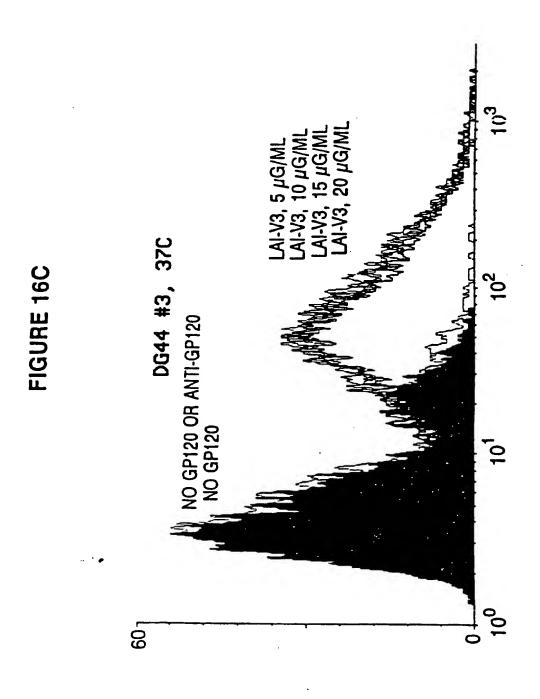


SUBSTITUTE SHEET (RULE 26)

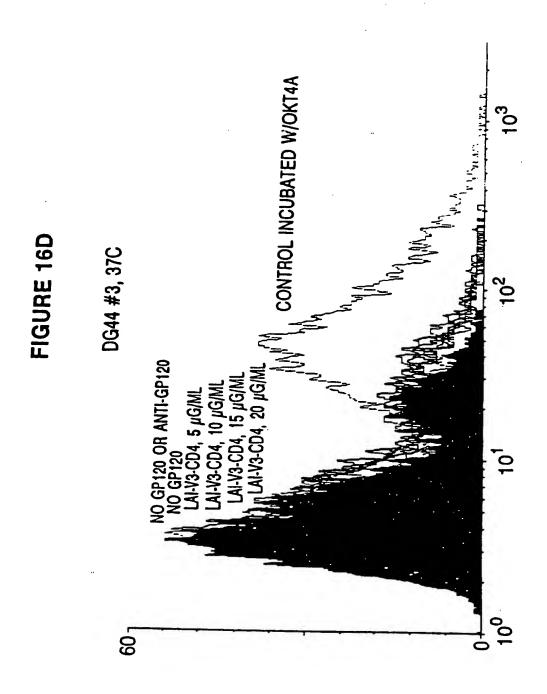


SUBSTITUTE SHEET (RULE 26)

41/42



SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)

#### INTERNATIONAL SEARCH REPORT

International application No. PCT/US94/03282

	والمتعالي والمراور والمراور والمناور والمتعالم والمتعالم والمناور والمتعالم والمتعالم والمتعارب والمراور		
A. CL	ASSIFICATION OF SUBJECT MATTER		
IPC(5)			
US CL	: 424/88, 89; 536/27; 530/395 to International Patent Classification (IPC) or to both	national classification and IPC	
1	LDS SEARCHED	of the stands of a single constants	
	documentation searched (classification system followed	od by cassification symbols)	
U.S. :	424/88, 89; 536/27; 530/395		
Document	tion searched other than minimum documentation to the	e extent that such documents are included in the fields searched	
		•	
Electronic	data base consulted during the international search (n	ame of data base and, where practicable, search terms used)	
APS. D	ialog, search terms: HIV-1, mutation, V3 loop,	C4 region, envelope glycoprotein, vaccines, nucleic acid	
C. DO	CUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where a	ppropriate, of the relevant passages Relevant to claim No.	
Y	Science, Volume 252, issued 1	7 May 1991, S. Wain- 6.7	
•	Hobson, et al, "LAV Revisited: C		
	Isolates from Institut Pasteur", p	<u> </u>	
	article.	- <b>G</b>	
		·	
Y	US, A, 5,030,449 (BERZOFSKY E	T AL) 09 July 1991, cols. 1-27	
·	4-7.		
	•		
Y	WO, A, 91/15512 (GREGORY E	T AL) 17 October 1991, 1-27	
`	entire patent.		
-	tituo patanti		
Y	WO, A, 91/11461 (PASEK ET AL) 08 August 1991, see 1-27 entire patent.		
`			
}			
İ			
ļ			
Furth	er documents are listed in the continuation of Box C	. See patent family annex.	
<u> </u>	cial categories of cited documents:	T later document published after the international filling date or priority	
•	ment defining the general state of the art which is not considered	date and not in conflict with the application but cited to underwand the	
to be of particular relevance  "X" document of particular relevance; the claimed invention cannot be			
extrest described power of canada to involve as investive sup			
document in taken sions  document in taken sions  cand to smallinh the publication date of another citation or other			
•	ind remains (as specified)	considered to involve an inventive step when the document is considered with one or more other such documents, such combination	
- doc	sonot referring to an oral disclosure, use, exhibition or other as	being obvious to a person skilled in the art	
	mens published prior to the international filing date but later than priority date claimed	"&" document member of the same patent family	
te of the a	ctual completion of the international search	Date of mailing of the international cearch report	
9 JUNE 1	004	JUN 2 4 1004	
7 JUNE 1	***		
	ailing address of the ISA/US er of Patents and Trademarks	Authorized officer W. Kyza for	
ox PCT		LYNETTE F. SMITH	
	D.C. 20231		
simile No	. (703) 305-3230 ) A/210 (second sheet)/July 1992)+	Telephone No. (703) 308-0196	
n FC 1/13	PULLU (SECURG ENCORPT) 1772/F		